

DISSERTATION

INFERENCE FOR CUMULATIVE INTRADAY RETURN CURVES

Submitted by

Ben Zheng

Department of Statistics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2018

Doctoral Committee:

Advisor: Piotr S. Kokoszka

Dan Cooley

Hong Miao

Wen Zhou

Copyright by Ben Zheng 2018

All Rights Reserved

ABSTRACT

INFERENCE FOR CUMULATIVE INTRADAY RETURN CURVES

The central theme of this dissertation is inference for cumulative intraday return (CIDR) curves computed from high frequency data. Such curves describe how the return on an investment evolves with time over a relatively short period. We introduce a functional factor model to investigate the dependence of cumulative return curves of individual assets on the market and other factors. We propose a new statistical test to determine whether this dependence is the same in two sample periods. The statistical power of the new test is validated by asymptotic theory and a simulation study. We apply this test to study the impact on individual stocks and Sector Exchanged-Traded Funds (ETF) of the recent financial crisis and of trends in the oil price. Our analysis reveals that the functional approach has an information content different from that obtained from scalar factor models for point-to-point returns. Motivated by the risk inherent in intraday investing, we propose several ways of quantifying extremal behavior of a time series of curves. A curve can be extreme if it has shape and/or magnitude much different than the bulk of observed curves. Our approach is at the nexus of Functional Data Analysis and Extreme Value Theory. The risk measures we propose allow us to assess probabilities of observing extreme curves not seen in a historical record. These measures complement risk measures based on point-to-point returns, but have different interpretation and information content. Using our approach, we study how the financial crisis of 2008 impacted the extreme behavior of intraday cumulative return curves. We discover different impacts on shares in important sectors of the US economy. The information our analysis provides is in some cases different from the conclusions based on the extreme value analysis of daily closing price returns. In a different direction, we investigate a large-scale multiple testing problem motivated by a biological study. We introduce mixed models to fit the longitudinal

data and incorporate a bootstrap method to construct an false discovery rate (FDR) controlling procedure. A simulation study is implemented to show its effectiveness.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Piotr Kokoszka, for his immeasurable guidance, support and encouragement during my program. He is such a good man of great character, and always provide me with wise advice about both scientific research and life in general. He has shown me, by his example, what a good researcher/person should be. Without him, I would have never been able to achieve so much and grow into a confident researcher.

I would also like to thank my doctoral committee members: Professors Dan Cooley, Hong Miao and Wen Zhou, for their valuable supports and interactions over the past few years. Each of them has provided me extensive personal and professional guidance and together formed a wonderful support system throughout my graduate career.

All of the projects in my dissertation are collaborations with other researchers. I thank Professor Stilian Stoev for his expertise in extreme value theory and the corresponding contributions to our research. I thank Professor Hong Miao for providing all relevant data sets and helping introduce our work to the finance community. I thank Professor Wen Zhou for his guidance in our research about the high dimensional testing problem.

I am grateful to many professors at Colorado State University who contributed to my knowledge in statistics and mentored me in teaching. I also want to thank my fellow graduate students and friends. Working together and becoming friends with you was such an amazing experience in my life, and this has helped contribute to my success here.

Lastly, many thanks to all my family members back in China for long-term and long-distance support. I thank my mother Lili Cao, my father Miao Zheng and my brother Zhou Zheng, your love are always with me and provide unending inspiration in whatever I pursue.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
Chapter 1 Introduction	1
Chapter 2 Two Sample Inference in a Factor Model for Cumulative Intraday Return Curves	7
2.1 Introduction	7
2.2 A Functional Factor Model	10
2.3 Two Sample Inference Test	12
2.3.1 Dependence Structure	12
2.3.2 The Test Statistic	13
2.3.3 Derivation of the Test Statistic and its Properties	16
2.4 Empirical Application	18
2.4.1 Data	19
2.4.2 Asymmetric CIDR Betas around the Financial Crisis	19
2.4.3 Asymmetric CIDR Betas for Bull and Bear Oil Price Periods	23
2.4.4 Comparison to Point to Point Betas	24
2.5 Finite sample performance	26
2.6 Concluding Remarks	29
2.7 Proofs	42
2.7.1 Proof of Theorem 5	42
2.7.2 Proof of Theorem 6	43
Chapter 3 Risk Analysis of Cumulative Intraday Return Curves	45
3.1 Introduction	45
3.2 Statistical Preliminaries	48
3.2.1 Background on Functional Data Analysis	48
3.2.2 Background on Extreme Value Theory	50
3.3 Risk analysis of CIDR curves	54
3.3.1 Extreme return curves	54
3.3.2 Data description	57
3.3.3 Extremes of the norm and the scores	58
3.3.4 Extremal dependence between the magnitude of the scores	64
3.3.5 Summary and concluding Remarks	67
3.4 Supplementary Materials	69
3.4.1 Variances of $\widehat{\text{VaR}}_\alpha$ and $\widehat{\text{ES}}_\alpha$	69
3.4.2 Estimation of the angular density	70
3.4.3 Elaboration on the measures χ and $\overline{\chi}$	71

3.4.4	Large tables and displays	78
Chapter 4	A Large-Scale Dynamic Cross-Correlation Testing Problem	96
4.1	Introduction	96
4.2	Motivation	97
4.2.1	Description of the data	97
4.2.2	The main objective	98
4.3	Preliminaries	98
4.3.1	A mixed model	99
4.3.2	Moving block bootstrap	101
4.4	Large-scale multiple testing procedure	103
4.4.1	FDR control procedure	104
4.4.2	FDR control procedure with MBB adjustment	107
4.5	Simulation	109
Chapter 5	Summary and Future Research	113
Bibliography	115

LIST OF TABLES

2.1	List of Tickers	30
2.2	Asymmetric Beta around the Financial Crisis (One Factor) – ETFs	31
2.3	Asymmetric Beta around the Financial Crisis (One Factor) – Stocks	32
2.4	Asymmetric Beta around the Financial Crisis (Two Factors) – ETFs	33
2.5	Asymmetric Beta around the Financial Crisis (Two Factors) -Stocks	34
2.6	Bull vs. Bear Oil Price Periods: ETFs	35
2.7	Bull vs. Bear Oil Price Periods: Individual Stocks	36
2.8	Two Sample Test Results – ETFs -Point to Point	37
2.9	Two sample test for the comparison two Factor Models -ETFs- Point to Point	38
2.10	Bull vs. Bear Oil Price Periods: ETFs -Point to Point	39
2.11	Monte Carlo Simulation Results of the test statistics $T_{M,N}$	40
3.1	Tickers of the assets (Sector ETFs) used in this study.	58
3.2	Time periods used in this study.	58
3.3	Estimation results for the norm $\ R\ $ for nine ETFs.	79
3.4	Estimation results for the magnitude of the first score $ \xi_1 $ for nine ETFs.	80
3.5	Estimation results for the magnitude of the point-to-point returns for nine ETFs.	81
3.6	Fisher’s exact test (with frequencies).	82
3.7	Fisher’s exact test (with probabilities).	83
3.8	Estimation results of extreme dependence between ξ_1^2 and ξ_2^2 for the nine ETFs.	84
4.1	Empirical FDR and Power for the LDCCT-B procedure.	111

LIST OF FIGURES

2.1	Ten Consecutive CIDs on JP Morgan Chase Stock. This figure shows ten consecutive days CIDR on the JP Morgan Chase Stock. While each curve is a realization of a non-stationary stochastic process, the sequence of curves can be assumed to be a realization of a stationary time series. The curves have approximately the same distribution as the Brownian motion.	41
3.1	Five Consecutive CIDs (sector ETF XLF). The noisy curves are the raw CIDs, the superimposed thick curves are the corresponding CIDs reconstructed from the first three FPCs. Details are explained in Section 3.2.	46
3.2	The first three FPCs for S&P500 during the period 2006/7/5 - 2011/12/30.	56
3.3	The scatter plot of $A_n^{(p)} = \{\sum_{j=1}^p \xi_{nj}^2\}^{1/2}$ against $\ R_n\ $ with $p = 2, 3, 4$ for XLF during the period 2008/6/2 - 2008/8/12 (50 days).	57
3.4	Left panel: histogram of the observed θ'_i s with the estimate $\hat{f}_h(\theta)$ superposed. Right panel: histogram of the same number of new θ'_i s generated from $\hat{f}_h(\theta)$	72
3.5	VaR & ES Estimates of norms of CIDs . This is an illustration of the VaR & ES Estimates (with standard errors) of norms of CIDs for the nine Sector EFTs.	85
3.6	VaR & ES Estimates of magnitude of point-to-point returns. This is an illustration of the VaR & ES Estimates (with standard errors) of magnitude of point-to-point returns for the seven Sector EFTs.	86
3.7	The scatter plot of XLF scores with VaR Estimates (Bootstrap). On each of the scatter plot, the big black thick circle corresponds to the norm at $\text{VaR}_{0.004}$ (or the 1-year return level), while the bigger grey circle corresponds to the norm at $\text{VaR}_{0.0001}$. The stars represent the real extreme points with norm greater than $\text{VaR}_{0.004}$. The little grey circles represent the simulated extreme points with norm greater than $\text{VaR}_{0.004}$ based on the Bootstrap method.	87
3.8	VaR Extremal regions of XLF (Bootstrap). This is the extremal regions of CIDs corresponding to extreme regions in different quadrants on the (ξ_1, ξ_2) plane in Figure 3.7. On each panel, the two curves in the same type, respectively, represents the lower and upper bound of the extremal region of CIDs regarding one of the four periods under comparison. The thick solid curves represent the real extremal CIDs with norm greater than $\text{VaR}_{0.004}$	88
3.9	The scatter plot of XLF scores with ES Estimates (Bootstrap). On each of the scatter plot, the big black thick circle corresponds to the norm at $\text{ES}_{0.004}$ (or the 1-year expected shortfall). The stars represent the real extreme points with norms greater than $\text{VaR}_{0.004}$. The little grey circles represent the simulated extreme points with norm at $\text{ES}_{0.004}$ based on the Bootstrap method.	89

3.10	ES Extremal regions of XLF (Bootstrap). This is the extremal regions of CIDRs corresponding to extreme regions in different quadrants on the (ξ_1, ξ_2) plane in Figure 3.9. On each panel, the two curves in the same type, respectively, represents the lower and upper bound of the extremal region of CIDRs regarding one of the four periods under comparison. The thick solid curves represent the real extremal CIDRs with norm greater than $\text{VaR}_{0.004}$	90
3.11	The scatter plot of XLF scores with VaR Estimates (KDE). On each of the scatter plot, the big black thick circle corresponds to the norm at $\text{VaR}_{0.004}$ (or the 1-year return level), while the bigger grey circle corresponds to the norm at $\text{VaR}_{0.0001}$. The stars represent the real extreme points with norm greater than $\text{VaR}_{0.004}$. The little grey circles represent the simulated extreme points with norm greater than $\text{VaR}_{0.004}$ based on KDE.	91
3.12	VaR Extremal regions of XLF (KDE). This is the extremal regions of CIDRs corresponding to extreme regions in different quadrants on the (ξ_1, ξ_2) plane in Figure 3.11. On each panel, the two curves in the same type, respectively, represents the lower and upper bound of the extremal region of CIDRs regarding one of the four periods under comparison. The thick solid curves represent the real extremal CIDRs with norm greater than $\text{VaR}_{0.004}$	92
3.13	The scatter plot of XLF scores with ES Estimates (KDE). On each of the scatter plot, the big black thick circle corresponds to the norm at $\text{ES}_{0.004}$ (or the 1-year expected shortfall). The stars represent the real extreme points with norms greater than $\text{VaR}_{0.004}$. The little grey circles represent the simulated extreme points with norm at $\text{ES}_{0.004}$ based on KDE.	93
3.14	ES Extremal regions of XLF (KDE). This is the extremal regions of CIDRs corresponding to extreme regions in different quadrants on the (ξ_1, ξ_2) plane in Figure 3.13. On each panel, the two curves in the same type, respectively, represents the lower and upper bound of the extremal region of CIDRs regarding one of the four periods under comparison. The thick solid curves represent the real extremal CIDRs with norm greater than $\text{VaR}_{0.004}$	94
3.15	Estimates of $\chi(q)$ and $\bar{\chi}(q)$ for the squared scores of XLF. This is the estimates results of $\chi(q)$ and $\bar{\chi}(q)$ for the squared scores of the first two CIDR FPCs for XLF. Dotted lines correspond to approximate 95% confidence intervals. From top to bottom, the panels show the results for the periods “before”, “during” “after1” and “after2”.	95

Chapter 1

Introduction

Functional data analysis (FDA) has established itself as an important field of statistics over the past two decades. It has found extensive applications in many areas such as biology, physics, geophysics, climate science, psychology, finance and economics. This field has become broad with many specialized directions of research. Functional time series, as a special subfield of functional data analysis, offers effective new tools in finance and economics research. In this dissertation, special attention is devoted to making inference for cumulative intraday return curves, introduced by [1] as a way of transforming the daily price curves into a stationary time series of functions.

This dissertation consists of three main parts. Chapters 2 and 3 are concerned with statistical modeling and inference for the commutative intraday return curves, which are the central themes of the whole dissertation, justifying the title. To illustrate the usefulness of our methodologies, a comprehensive application is made to common financial assets, especially to their behavior with respect to the 2008 financial crisis. Chapter 4 presents a stand alone study of a large scale cross correlation testing problem motivated by a biological experimental study. Before diving into details of each individual research topics, we describe in the rest of this chapter some relevant concepts as a preliminary, and give general introductions to the main ideas of each chapters, as well as highlighting some useful findings in each study.

Many financial data sets form functional time series. Cumulative intraday return curves (CIDRs) are data of this form that can be used to study dynamic performance of financial assets. Suppose P_n represent the closing price of a given financial asset on day n . Then the daily closing return is defined as

$$r_n = 100(\ln P_n - \ln P_{n-1}).$$

This is exactly the definition of the point-to-point return which is commonly used in conventional financial models such as the Capital Asset Pricing Model (CAPM). Now suppose $P_n(t)$ is the price

of a financial asset at time t within day n . Then the CIDRs are defined as functions

$$R_n(t) = 100(\ln P_n(t) - \ln P_n(0)).$$

One can clearly see that the CIDR is a counterpart of the point-to-point return, but instead of selecting one point to represent the return over a whole day, the CIDR contains some information about the evolution of the return within a day. It may be expected to provide additional information contained in the function that is not available from application of traditional methods. As we will see, inference based upon CIDRs is often more robust than its scalar counterpart, because it averages out random influence on each single point and focuses on the overall influence of the daily curves as a whole. Another main strength of CIDRs is that, in contrast to traditional high-frequency data analysis, there is no concern about serial dependence among repeated measurements because we treat the whole CIDR curve as a single statistical entity. The aforementioned merits inspire us to use CIDRs as a new tool to reexamine many existing results or conclusions in financial research. At the theoretical level, we rigorously define CIDRs as a special example of the functional time series, and develop theory grounded in the Hilbert space formalism. For more rigorous definition and basic assumptions imposed upon CIDRs, the reader is referred to Section 2.3.

The first important problem in finance we examined under the new functional data framework is the study of asymmetric correlations/betas. In finance, beta is a term that tells us how the expected return on an asset changes as the expected market returns change. It is a slope of a specific regression line. For example, assume the S&P 500 is the index summarizing “the market” and a stock’s beta is 1.2, it means, theoretically, the stock is 20% more volatile than “the market”. So understanding how the betas change is very useful for understanding the riskiness of the underlying assets. Correlation is another important measure closely related to beta. The term “asymmetric correlations” refers to imbalanced correlations of an asset during market downturns and upturns. More details about this issue are provided in Section 2.1. To study this asymmetric beta/correlation problem with CIDRs, we need to bring in a functional regression model, which extends the classical capital asset pricing model (CAPM). In Chapter 2 we introduce a functional regression model

with CIDRs of individual asset as the responses and CIDRs of the market and/or other factors as predictors. The regression coefficients, or betas for consistency with the scalar context, are interpreted in a similar way to their point-to-point return based counterparts. We use the mathematical formalism developed by [2], who proposed a way of estimating betas in a functional regression model by combining least square and the method of moments approaches. Our major challenge is to find a way of comparing betas estimated from two independent samples. A main contribution of our work is to propose a statistical two-sample test associated with a functional regression model. The testing procedure has an elegant and simple form, and only a few weak assumptions are needed. The null distribution of the test statistic is asymptotically Chi-square with a known number of the degrees of freedom. There are at least twofold contribution of this work. Even though it is motivated by a specific question of asymmetric correlation of CIDRs, the test procedure can be applied to comparing betas for any stationary functional time series based factor model given all the assumptions are satisfied. On the other hand, the conventional scalar tests are essentially a special case of our test, so, we offer a unified framework to study the asymmetric correlation problems. As an illustration of the proposed methodology, an extensive application is made to a selection of financial assets such as individual US stocks and Sector ETFs (Exchange-Traded Funds) of the periods before, during and after the 2008 crisis. Details of the data analysis are well documented in tables in Chapter 2 along with elaborations on the findings.

As noted above, CIDRs are essentially normalized/standardized intraday price curves. A CIDR curve has both shape and magnitude which contain much more abundant information than a point-to-point return. For example, the shape can reflect how a stock return evolves over time in a day while the magnitude reflects the overall daily volatility/momentum of the stock return. Dealing with CIDRs represents a change in philosophy towards the handling of financial data. The primary objective in Chapter 3 is to quantify extremal behaviors of a time series of curves in the context of CIDRs. Imagine you are a quantitative researcher investing in intraday prices and the valuations of your position might be impacted by the shape of the daily price curves. It will be useful for you to know the probability of observing some unusual shapes ahead of time. One of the major

accomplishments we achieved in Chapter 3 is to develop a procedure of creating a region for extreme curves based on past data. Back to the day trader example, suppose one day you observe a CIDR curve falling in the predicted extremal region with an upper tail probability of 0.001, you may be aware of that this is a rare curve shape that should be only observed once every four years.

In Chapter 3, we rely on two important statistical techniques to achieve our goal of quantifying extremal curves. The first one, also a core technique in functional data analysis, is called Functional Principal Component Analysis (FPCA). It is an important dimension reduction tool that can be used to investigate the dominant modes of variation of functional data. Specifically, by choosing a set of orthonormal eigenfunctions as the eigenbasis in a Hilbert space, a random function can be represented in the most parsimonious way, that is, one can satisfactorily approximate a random function which lives in an infinite dimensional space using a small number of coefficients. In applications, without losing much information, this is an efficient way of transforming a functional data problem into a multivariate vector problem. The reader is referred to Section 3.2.1 for more details about definitions and properties of FPCA related concepts. In our study, FPCA reveals some interesting findings about the shape of the CIDRs. By using functional principal component analysis, we can tell what types of shape form/dominate the daily variation of an intraday price curve. We note here the recent approach of [3] who develop principal components specifically tailored to the study of extremes. Their approach may also be useful in the study of risk. The other important statistical tool used in this study is called peak-over-threshold (PoT), which is one of the two related methods in the extreme value theory (EVT). The basic idea of PoT is to model the distribution of exceedances over a certain threshold as a Generalized Pareto Distribution (GPD). The other major EVT method is called block-maxima-method (BMM). This method models the distribution of the maxima of each pre-specified block (or period), which is a more traditional way of identifying extremes in seasonal data. However, [4] argue that the PoT method is superior in financial applications. In Chapter 3, we choose the PoT method as a better fit for our data for a similar reason after a preliminary exploratory study of the data. Another strength of using GPD instead of nonparametric methods is that it allows us to depict possible shapes of extreme curves

that may never be observed in sample curves. The reader is referred to Section 3.2.2 for a detailed summary of PoT method. Again, a practical application is presented to nine major Sector ETF CIDRs. In particular, extreme regions of specified tail risk measures are created for one index as an illustration of the proposed method.

Lastly, Chapter 4 addresses an unrelated large scale multiple testing problem. Problems of this type originate from genomic and genetic research, where usually tens of thousands of individual hypotheses are tested simultaneously. A major difficulty comes from a simple fact that when the number of individual tests becomes large, it will be very hard to control the overall family-wise error rate at a certain level as well as maintaining a good power. In most of the recent research on this topic, the focus is on controlling the false discover rate (FDR), which is defined to be the expected proportions of false positives among all the rejections. The objective of our project is to simultaneously test the hypotheses

$$H_{0pg} : \lambda_{pg} = 0 \quad \text{versus} \quad H_{1pg} : \lambda_{pg} \neq 0 \quad \text{for } 1 \leq p \leq P, 1 \leq g \leq G,$$

where λ_{pg} represents the cross correlation between g th gene and p th protein. This is a typical large scale cross correlation testing problem as the number of genes and proteins G and P are large while the sample size is rather small. For details of the special data structure, see Section 4.2.1. A contribution most closely to Chapter 4 is [5] who developed a new scheme for large scale multiple testing of cross correlations with the assumption of the sparse true significant alternatives. On one hand, in order to fit this idea in the longitudinal-like data structure, we use a mixed model based coefficient instead of the standard correlation coefficient to construct the individual test statistic. A specialized bootstrap procedure is utilized to estimate each marginal null distributions for the sake of the longitudinal-like data structure and small sample size. Finally, an algorithm based on the newly proposed testing procedure is established and a simulation study is presented to show its effectiveness. According to the simulation results display in tables, the major flaw in the proposed method is its reliance on the assumption of sample size being sufficient big. Unfortunately, this assumption is not true for the data we have at present.

Chapter 2 has been published as the following paper “Kokoszka, P., Miao, H. and Zheng, B. (2017). Testing for asymmetry in betas of cumulative returns: Impact of the financial crisis and crude oil price. *Statistics & Risk Modeling*, 34.1-2:33-53". Chapter 3 has been accepted as the following paper “Kokoszka, P., Miao, H., Stoev, S. and Zheng, B. (2018). Risk analysis of cumulative intraday return curves. *Journal of Time Series Econometrics*". Chapter 4 is a third joint project with authors Wen Zhou, Piotr Kokoszka and Ben Zheng, which remains in progress.

Chapter 2

Two Sample Inference in a Factor Model for Cumulative Intraday Return Curves

2.1 Introduction

The contribution of this paper is twofold. At the methodological level, we propose a statistical inferential tool to test the equality of functional regression betas between two samples. At the empirical level, we utilize the new tool to investigate the impact of the 2007 to 2009 financial crisis and the regime of increasing or decreasing oil prices on suitably defined betas of intraday cumulative returns. Our analysis uses cumulative intraday return (CIDR) curves computed from high frequency data. Such curves describe how the return on an investment evolves with time over a relatively short period.

The study of asymmetric correlations, a term referring to different correlations between security returns during market downturns and upturns, is important for at least two reasons. First, hedge ratios rely crucially on the correlations between the assets to be hedged and the associated financial instruments. The presence of asymmetric correlations can impact hedging effectiveness. Asymmetric correlations require dynamic changing of hedge ratios conditional on market conditions. Second, portfolio diversification is the center piece of standard investment theory. Studies find that correlations between stocks tend to be much higher as the market falls. Thus, the effectiveness of diversification might be questionable, particularly during bear markets. Asymmetric correlations of point-to-point stock returns with market indexes are well documented in the finance literature.

Beta is the product of the correlation and the volatility of an asset relative to the volatility of the benchmark (the market). Therefore, asymmetric correlations and asymmetric beta are similar. Betas are closely related to asset pricing theories, and useful in understanding the riskiness of the associated assets.

Using three alternative definitions of bull and bear markets, [6] find that the betas were not significantly different in the two types of markets for most of the 700 stocks studied (all the stocks traded at NYSE at that time). [7] show that most of the tested 322 securities exhibited betas which were not significantly different in up and down markets. [8] revisit the [6] procedures for new samples. Unlike the findings of [6], they find statistically different betas for stocks over bull and bear markets. More recently, [9] use a logistic smooth transition market (LSTM) model to investigate whether ‘bull’ and ‘bear’ market betas for Australian industry portfolios differ. Their results indicate that ‘bull’ and ‘bear’ betas are significantly different for most industries, and that up-market risk is not always lower than down-market risk. Other studies discover asymmetric betas on portfolios based on different criteria, such as, size based portfolios ([10], [11] and [12]), risk based portfolios ([13] and [11]) and past performance based portfolios ([11] and [14]).

[15] find strong asymmetric correlations between stock portfolios and the US market. [16], [17], [18] and [19], among others, document asymmetries in betas of stock returns, but provide no formal statistical tests. [20] develop a model-free test for asymmetric correlations and provide formal tests for asymmetric betas and covariances. They sort stocks by size, book-to-market, and momentum and find strong evidence of asymmetries for both size and momentum portfolios, but no evidence for book-to-market portfolios. They also show that incorporating asymmetries into investment decisions can be of substantial importance for an investor with disappointment aversion. All the above studies use monthly closing price return data.

While point-to-point returns using closing monthly or daily prices, have been extensively studied, there is no work we are aware of that pertains to cumulative intraday returns. While two sample problems are among the most thoroughly studied of statistical issues, there is relatively little research that examine samples of curves, and none in the context of functional regression analysis. Functional regression analysis is the regression of cumulative return curves for individual assets against cumulative return curves of the market and/or other factors.

Although two sample problems are considered by [21], [22], [23], [24] and [25], two sample inference in the context of functional regression models has not been previously studied.

We propose a statistical test in the context of functional data analysis. The test can be broadly used to examine the statistical equality of the regression coefficients of the same model estimated from two samples. Under suitable asymptotic assumptions, the properties of the statistical test are established and proved. Then the statistical test is used to investigate whether the betas of individual assets are different before, during and after the financial crisis. The 30 components of the Dow Jones Industrial Average and the 9 Selected Sector ETFs are the assets used in the study. The results show that roughly one third of the individual stocks have statistically significantly different market betas before and during (10 stocks), during and after (11 stocks) and before and after (12 stocks). For the ETFs, only one has a different beta before and during the crisis, five have different betas during and after, and five have different betas before and after the financial crisis.

To further show the empirical power of the test, a crude oil price factor is added to the functional regression model. Studies in the finance literature have shown that crude oil prices is an important factor that impacts the economy as well as the equity markets. [26] examined the correlation between the price of crude petroleum and U.S. recessions. He finds that oil shocks were a contributing factor in at least some of the U.S. recessions prior to 1972. [27] use change in the monthly average U.S. dollar price per barrel of crude oil as a factor for cross-sectional expected returns. [28] document that in the postwar period, the reaction of the United States and Canadian stock prices to oil shocks can be completely accounted for by the impact of these shocks on real cash flows. After the financialization of commodity futures markets in 2004–05, oil volatility has become a strong predictor of returns and volatility of the overall stock market. Furthermore, stocks' exposure to oil volatility risk now drive the cross-section of expected returns. [29] investigated the impact of daily oil price changes on the stock returns of a wide array of industries and found that in addition to the stock returns of industries that depend heavily on oil, stock returns of some industries that use little oil are also sensitive to oil prices perhaps because their main customers are impacted by oil price changes.

First, the impact of the crude oil factor is studied by using it as an additional factor in the comparison of the before, during and after financial crisis functional factor models. The results

show that when the crude oil factor is added to the models, the test shows that betas for five, seven and nine out of the nine sector ETFs are different for the three paired samples, respectively. For the individual stocks, the number of significantly different betas are also dramatically increased for all of the three paired subsamples. Since the test is a joint test, and in the above mentioned results, we already tested the significance of the asymmetric market beta, the increases in the number of significantly different betas are due to the impact of the crude oil factor.

Subsequently, different samples conditional on the trend of the crude oil prices are used to investigate the impact of the crude oil factor from a different angle. The test results show that during “bull” and “bear” crude oil prices, the betas for ETFs are different for five out of the nine sectors. For the individual stocks, the results show that the betas for most of the stocks are insignificant during the “bull” oil price subsample, and positive and significant during the “bear” oil price subsample. That is, negative crude oil price movements have a negative impact on intraday cumulative equity returns.

Furthermore, a comparison is made of the results of the functional factor model with the results of the point-to-point return model. The results show that the functional factor models which considers intraday behavior, or in other words how the price moves within trading days, reveal new information content regarding the relationships between equity price movement and the factors.

The remainder of the chapter is organized as follows. Section 2.2 introduces the model structure. After a general background has been established, Section 2.3 describes the test procedure and the properties of the test. Section 2.4 presents its application to cumulative intraday returns on blue chip stocks and sector ETF’s, and a comparison of the findings to those obtained using point-to-point returns. Section 2.5 utilizes Monte Carlo simulations to examine the empirical size and power of the test and Section 2.6 concludes the paper and proposes future research directions.

2.2 A Functional Factor Model

We study the behavior of a time series of financial asset returns from a functional data analysis perspective. We view the series of Cumulative Intra-Day Returns (CIDRs) as a time series of

functions. Since this concept has not been widely used in the finance literature, its definition is now shown.

Suppose $P(t_0)$ is the price of an asset at time t_0 , a conveniently selected time within a trading day. For instance, t_0 can be the NYSE opening time, i.e. 9:30am EST. The cumulative intraday return is defined as $R(t) = \log P(t) - \log P(t_0)$, $t_0 < t \leq t_0 + T$, where, T is number of sample unit time (i.e. minute, second) during a trading day. Thus, the object of the study is a set of Cumulative Intraday Returns (CIDRs) defined by

$$R_n(t) = \log P_n(t) - \log P_n(t_0), \quad t_0 < t \leq t_0 + T, \quad (2.1)$$

where n indexes the trading day, t_0 is the opening time of the trading day, and T is the number of time units in a trading day. One minute is used as the sample frequency, so $T = 390$ minutes for a typical NYSE trading day. The CIDR curves $R_n(t)$ have shapes almost identical to the intraday price curves. This follows from the approximation that $R_n(t) \approx (P_n(t) - P_n(t_0))/P_n(t_0)$ and the observation that for any day n , $P_n(t_0)$ is a fixed number. The curves $R_n(t)$ start from zero. Even though for each n , the curve $R_n(t)$ is a realization of a nonstationary stochastic process, basically a Brownian motion, the sequence of curves is stationary in a function space (c.f. [30]). This is the main reason for studying the CIDR's instead of the price curves. At the methodological level, this allows us to focus on the intraday behavior of price curves.

To investigate the relationship between the cumulative intraday return curves of an asset and some factors, such as the intraday cumulative return curves of the market, [2] postulate the following model:

$$R_n(t) = \beta_0(t) + \sum_{j=1}^p \beta_j F_{nj}(t) + \varepsilon_n(t), \quad (2.2)$$

in which β_0 is an asset dependent intercept curve, F_{nj} is the CIDR curve of factor j , (e.g. CIDR on a market index or oil price), and the ε_n are error curves. This model is analogous to the commonly used factor models, i.e. the Capital Asset Pricing Model (CAPM), but the independent and the dependent variables are curves instead of scalar time series. The focus of this work is on the

coefficients β_j which describe the dependence of the CIDR's of an individual asset on the factors F_{nj} . The rationale for such a focus is that the intraday price curves of individual assets are generally proportional to the market index price curves. In the spirit of the CAPM, this proportionality should be even more pronounced when considered in the context of cumulative returns. In the application, the interest is in whether the betas β_j in (2.2) are different over different periods. This is not trivial since the dependent and independent variables are curves, and the standard t-test does not apply. Thus, in the next section, a two sample statistical test is developed applicable to the new functional settings.

2.3 Two Sample Inference Test

Two periods of time are considered, the first period includes N trading days, the second M trading days. The two samples are assumed to be independent. However, since each sample generally consists of data on consecutive trading days, serial independence within each sample is not assumed. Instead, a general nonparametric form of dependence is assumed.

2.3.1 Dependence Structure

We assume a very general form of dependence which has been extensively employed in time series research for over a decade. It relies on the representation of a stationary time series in an abstract space as a Bernoulli shift of independent and identically distributed (*i.i.d.*) errors. See Chapter 16 of [31] for historical background, detailed discussion and additional references. Our results could also be proven using different dependence notions.

Suppose H is a separable Hilbert space, as shown by [32]. Let $p \geq 1$ and let L_H^p be the space of H -valued random elements X such that

$$\nu_p(X) = (E\|X\|^p)^{1/p} < \infty.$$

In our application, the Hilbert space in question is the space of square integrable functions on the interval $[0, 1]$, and so the norm is defined by $\|X\|^2 = \int_0^1 X^2(t)dt$.

Definition 1. A sequence $\{X_n\} \in L_H^p$ is called L^p - m -approximable if each X_n admits the representation

$$X_n = f(u_n, u_{n-1}, \dots), \quad (2.3)$$

where the u_i are *i.i.d.* elements taking values in a measurable space S , and f is a measurable function $f : S^\infty \rightarrow H$. Moreover, we assume that if $\{u'_i\}$ is an independent copy of $\{u_i\}$ defined on the same probability space, then letting

$$X_n^{(m)} = f(u_n, u_{n-1}, \dots, u_{n-m+1}, u'_{n-m}, u'_{n-m-1}, \dots) \quad (2.4)$$

we have

$$\sum_{m=1}^{\infty} \nu_p(X_n - X_n^{(m)}) < \infty. \quad (2.5)$$

The gist of Definition 1 is that the dependence of f in (2.3) on the innovations far in the past decays so fast that these innovations can be replaced by their independent copies. Such a replacement is asymptotically negligible in the sense quantified by (2.5). The reader will notice that if $f(u_n, u_{n-1}, \dots) = \sum_{j=0}^{\infty} c_j u_{n-j}$, a dependent linear model is recovered. A nonlinear f allows a fairly general nonlinear form of dependence.

2.3.2 The Test Statistic

For simplicity, we use the superscript “*” to denote the observations taken in the second period. Thus, the first sample follows model (2.2), and the second sample follows the model

$$R_m^*(t) = \beta_0^*(t) + \sum_{j=1}^p \beta_j^* F_{mj}^*(t) + \varepsilon_m^*(t), \quad m = 1, 2, \dots, M. \quad (2.6)$$

We wish to test the null hypothesis $H_0 : \beta = \beta^*$, where

$$\beta = [\beta_1, \dots, \beta_p]^T, \quad \beta^* = [\beta_1^*, \dots, \beta_p^*]^T.$$

We want to determine whether the vectors of the coefficients of the factors are different over the two sample periods. In the simplest case of a single factor, which is the CIDR on a market index, this corresponds to checking whether the asset betas are statistically the same. We emphasize that even in this case our approach provides information which has not been available so far since we study the dependence of the shape of the intraday price curve of an asset on the shape of the intraday price curve of a market index, whereas previous research only considers point-to-point returns. In other words, we care about how and how much the price moves whereas the traditional model only concerns how much the price changes.

The test is based on the comparison of suitable estimators $\hat{\beta}$ and $\hat{\beta}^*$. We now define the estimator $\hat{\beta}$ introduced by [2], $\hat{\beta}^*$ is defined analogously. To simplify formulas, all functions (e.g. R_n, F_{nj}) appearing below are assumed to be defined on the unit interval $(0, 1]$. Thus, before performing the test, $t \in (t_0, t_0 + T]$ must be replaced by $(t - t_0)/T$. In other words, the CIDR functions are rescaled to the unit interval. To further lighten the notation, we introduce the inner product

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt. \quad (2.7)$$

With this notation, we define

$$\hat{\mathbf{F}} = \left[N^{-1} \sum_{n=1}^N \langle F_{nj} - \bar{F}_j, F_{nk} - \bar{F}_k \rangle, j, k = 1, 2, \dots, p \right] \quad (p \times p), \quad (2.8)$$

$$\hat{\mathbf{R}} = \left[N^{-1} \sum_{n=1}^N \langle R_n - \bar{R}, F_{nj} - \bar{F}_j \rangle, j = 1, 2, \dots, p \right]^T \quad (p \times 1), \quad (2.9)$$

where $\bar{R}(t) = N^{-1} \sum_{n=1}^N R_n(t)$ and $\bar{F}_j(t) = N^{-1} \sum_{n=1}^N F_{nj}(t)$, ($j = 1, \dots, p$). The matrix $\hat{\mathbf{F}}$ includes cross-sectional sample covariances of the factor curves, the vector $\hat{\mathbf{R}}$ consists of the sample covariances of the CIDRs on the asset of interest with the factor CIDRs. The estimator of [2] is defined as

$$\hat{\beta} = \hat{\mathbf{F}}^{-1} \hat{\mathbf{R}}. \quad (2.10)$$

It is a hybrid method of moments/least squares estimator which is consistent and asymptotically normal. The intercept function is estimated by

$$\hat{\beta}_0(t) = \bar{R}(t) - \sum_{j=1}^p \hat{\beta}_j \bar{F}_j(t). \quad (2.11)$$

A natural tests statistic would be $\|\hat{\beta} - \hat{\beta}^*\|^2 = \sum_{j=1}^p (\hat{\beta}_j - \hat{\beta}_j^*)^2$, but its asymptotic distribution is very complex, and even a Monte Carlo test based on it would be extremely time consuming to implement. However, a quadratic form of $\hat{\beta} - \hat{\beta}^*$ can be constructed whose limit distribution, as $N, M \rightarrow \infty$, is the standard chi-square distribution with p degrees of freedom. To define this quadratic form, we first introduce the long-run covariance matrices which describe serial dependence within each sample. We again focus on the first sample, with the understanding that everything is defined analogously for the second sample.

For $j = 1, \dots, p$, set $\mu_j(t) = EF_{nj}(t)$, and introduce the vectors

$$\xi_n = [\xi_{n1}, \dots, \xi_{np}]^T = [\langle \varepsilon_n, F_{n1} - \mu_1 \rangle, \dots, \langle \varepsilon_n, F_{np} - \mu_p \rangle]^T. \quad (2.12)$$

Denote by Γ their long-run covariance matrix defined by

$$\Gamma = \sum_{h=-\infty}^{\infty} E[\xi_0 \xi_h^T]. \quad (2.13)$$

Let $\hat{\Gamma}$ be a consistent estimator of Γ . Many estimators are implemented in statistical packages. In our empirical application, we used the R function `lrvar` (with default tuning parameters) in the package `sandwich`. An estimator must be applied to sample versions of the unobservable vectors ξ_n . The j th component of $\hat{\xi}_n$ is defined by

$$\hat{\xi}_{nj} = \langle \hat{\varepsilon}_n, F_{nj} - \bar{F}_j \rangle, \quad (2.14)$$

$$\hat{\varepsilon}_n(t) = R_n(t) - \hat{\beta}_0(t) - \sum_{j=1}^p \hat{\beta}_j F_{nj}(t). \quad (2.15)$$

Once estimators $\hat{\Gamma}$ and $\hat{\Gamma}^*$ have been computed, we can calculate the matrix

$$\hat{D}_{M,N} := \frac{N+M}{N} \hat{\mathbf{F}}^{-1} \hat{\Gamma} \hat{\mathbf{F}}^{-1} + \frac{N+M}{M} (\hat{\mathbf{F}}^*)^{-1} \hat{\Gamma}^* (\hat{\mathbf{F}}^*)^{-1}. \quad (2.16)$$

The test statistic is then defined as the quadratic form:

$$T_{M,N} := (N+M)(\hat{\beta} - \hat{\beta}^*)^T (\hat{D}_{M,N})^{-1} (\hat{\beta} - \hat{\beta}^*). \quad (2.17)$$

2.3.3 Derivation of the Test Statistic and its Properties

The previous subsection established a quantity to test the hypothesis. This subsection provides the properties of the test. For simplicity, we only state the assumptions and results for one sample, with the understanding that the same assumptions and results hold for the second sample with index m and superscript “*” added. Assumption 2 states that the factor and error functions are general nonlinear moving averages of some abstract errors. Assumption 3 states usual moment and independence assumptions in the context of random functions.

Assumption 2. Suppose there exist independent sequences $\{\delta_n\}$ and $\{\eta_n\}$ consisting of *i.i.d.* random variables taking values in measurable spaces \mathcal{S}_δ and \mathcal{S}_η , respectively. Assume there are measurable functions f_j and e defined on the appropriate product spaces, such that

$$F_{nj} = f_j(\delta_n, \delta_{n-1}, \dots), \quad \varepsilon_n = e(\eta_n, \eta_{n-1}, \dots).$$

Assumption 3. The factors $\{F_{nj}, j = 1, \dots, p\}$ and the errors $\{\varepsilon_n\}$ are random elements of the space $L^2([0, 1])$ which satisfy $E \|F_{nj}\|^4 < \infty$, $E \|\varepsilon_n\|^4 < \infty$, and $E \varepsilon_n = 0$. Moreover, for each n , the error functions $\{\varepsilon_n\}$ are independent of the vector factors $\{\mathbf{F}_n\}$.

Under Assumptions 2 and 3, letting $\mathbf{F}_n = [F_{n1}, F_{n2}, \dots, F_{np}]^T$, it follows that the vector factors $\{\mathbf{F}_n\}$ and the errors $\{\varepsilon_n\}$ are stationary random elements of the spaces, respectively, $(L^2)^p$ and L^2 .

The above assumptions must hold for both samples. Intuitively, these assumptions require that the CIDR's R_n and the factors F_{nj} form stationary sequences of functions (for each sample separately). As noted in the Introduction, the above stationarity has been verified using the tests of [30] which are implemented in the R package `fts.a`. Examination of the graphs, c.f. Figure 2.1, provides a visual justification.

Our last assumption requires the sample sizes to be comparable, that is:

Assumption 4. Assume that as $N, M \rightarrow \infty$,

$$\frac{N}{N+M} \rightarrow \theta, \quad \text{for some } 0 < \theta < 1. \quad (2.18)$$

The following Lemma, which follows from Theorem 1.2 in [2], provides the first step in the derivation of the asymptotic distribution of the test statistic $T_{N,M}$.

Lemma 1. Suppose Assumptions 2 and 3 hold, the sequences $\{\varepsilon_n\}$ and $\{\mathbf{F}_n\}$ are L^4 - m -approximable, and the matrix \mathbf{F} defined by

$$\mathbf{F} = [E \langle F_{nj} - \mu_j, F_{nk} - \mu_k \rangle, j, k = 1, 2, \dots, p] \quad (p \times p). \quad (2.19)$$

is nonsingular. Then,

$$\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} \mathbf{F}^{-1} \mathbf{W} \quad (2.20)$$

where \mathbf{W} is a p -dimensional mean zero Gaussian distribution with covariances $\text{Var}(\mathbf{W}) = \mathbf{\Gamma}$, and $\mathbf{\Gamma}$ is the long run covariance matrix defined by (2.13).

Relations leading to (2.20) also hold for the second sample. This is achieved by replacing all n by m and adding the superscript $*$. We thus have

$$\sqrt{M}(\hat{\beta}^* - \beta^*) \xrightarrow{d} (\mathbf{F}^*)^{-1} \mathbf{W}^*, \quad (2.21)$$

where \mathbf{W}^* is a p -dimensional mean zero Gaussian distribution with covariances $\mathbf{\Gamma}^*$.

The next theorem establishes the asymptotic distribution of $T_{N,M}$ under H_0 .

Theorem 5. Suppose that Assumptions 2 to 4 hold, the sequences $\{\varepsilon_n\}$ and $\{\mathbf{F}_n\}$ are L^4 - m -approximable, the matrix (2.19) is nonsingular, and the long-run covariance matrix $\mathbf{\Gamma}$ is nonsingular and is consistently estimated by $\widehat{\mathbf{\Gamma}}$. Also assume that the second sample satisfies analogous assumptions. Then, under H_0 ,

$$T_{N,M} := (N + M)(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}^*)^T (\widehat{\mathbf{D}}_{M,N})^{-1} (\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}^*) \xrightarrow{d} \chi^2(p). \quad (2.22)$$

Proof. See Section 2.7.1. □

Theorem 5 states that under suitable assumptions the asymptotic distribution of the statistic $T_{M,N}$ converges to the $\chi^2(p)$ distribution, if H_0 is true. We now establish the consistency of the testing procedure.

Theorem 6. Suppose the Assumptions of Theorem 5 hold, and $\boldsymbol{\beta}^* = \boldsymbol{\beta} + \boldsymbol{\eta}$ for some nonzero vector $\boldsymbol{\eta}$. Then, $T_{N,M} \xrightarrow{P} \infty$.

Proof. See Section 2.7.2. □

Theorem 6 implies that the power of the test approaches unity, as $N, M \rightarrow \infty$. That is, for large N and M , the probability of rejection of the null hypothesis if alternative is true converges to 100%.

2.4 Empirical Application

In this section, we apply the test introduced in Section 2.3 in several settings. First the data, the assets, and factors are summarized. Next, Subsection 2.4.2 presents the results of the application of our test to periods before, during and after the financial crisis. Subsection 2.4.3 reports a similar analysis for the “bull” and “bear” crude oil price periods. Finally, in subsection 2.4.4, the results are compared with those obtained by applying a simplified form of our test to point-to-point returns

on the same assets and over the same time periods to compare the information contents of the cumulative return curves and the point-to-point returns.

2.4.1 Data

We are interested in examining whether the betas for individual stocks and “portfolios” are time-varying. We also want to show that our approach extracts different information than the classic factor models. Thus, our approach is applied to the 30 stocks from the Dow Jones Industrial Average Components during the study period. As proxies for “portfolios”, the nine Select Sector SPDR ETFs are used. The nine Select Sector SPDR ETFs are Exchange Traded Funds that tracks the nine S&P 500 sector indexes. The ETFs hold individual stocks within the corresponding sectors. For ease of reference, Table 2.1 lists the companies and the Select Sector SPDR ETFs.

The factors include the market factors and the crude oil factor. The Dow Jones Industrial Average (DJ) is used as the market factor for individual stocks, and the S&P 500 Index (SP) as a proxy for the market for the ETFs. The crude oil factor is represented by the nearest term crude oil futures contract (CL). To construct continuous time series, we roll the nearest contracts to the next contract 20 calendar days before the expiration of the contracts (roughly at the beginning of each month). The Crude Oil Futures contracts are traded at the CME group.

The one minute frequency stock and ETF price data is obtained from Quantquote whereas the price data for the indexes and futures is from TickData.

2.4.2 Asymmetric CIDR Betas around the Financial Crisis

In this section, we apply the statistical test to responses and factors listed in Section 2.4.1 around the recent financial crisis. Our objective is to determine whether there are pair-wise statistically significant differences for betas during periods before and during, during and after, and before and after the financial crisis, respectively. Comparing betas for periods around the financial crisis is particularly interesting, as it allows us to determine whether the crisis has changed the relationships between the factors and individual assets reflected by the betas defined in this paper. The periods are defined as: “Before” (06/13/2006 – 09/27/2007), “During” (11/01/2007 –

02/27/2009), and “After” (01/04/2010 – 12/30/2011) the financial crisis. The number of sample days for the three periods are 326, 328 and 503 days, respectively. The sample periods are selected after a preliminary analysis.¹

We first run the one factor model and apply the test to the models for the three periods. The model is of the following form:

$$R_n(t) = \beta_0(t) + \beta_1 M_n(t) + \varepsilon_n(t), \quad (2.23)$$

where $R_n(t)$ is the CIDR (a function of t) on day n , $M_n(t)$ is the CIDR of the S&P 500 index for the ETFs, and the CIDR of the Dow Jones Industrial Average Index for individual stock.

Table 2.2 presents the estimated results of the model of Equation (2.23) for the three periods. The Table includes the estimated betas, the standard error of the beta estimates and the P-values of the pair-wise difference tests using the above developed statistic test. The results show that all the betas are highly significantly different from zero as indicated by the fact that the ratios of the estimators and the corresponding errors are between 5.00 (XLP for the “During” period) and 16.25 (XLE for the “After” period). In general the curve betas are consistent with common sense. For instance, for all the three periods, the economically sensitive sectors of XLE (energy), XLB (materials), XLI (industrials), XLF (financials), and XLK (technology) have higher betas for all periods than the other sectors.

The P-values show that for eight out of the nine sector ETFs, only XLK has significantly different betas before and during the financial crisis (0.86 vs 0.61). For the other ETFs, although the estimated betas are much higher for XLF during the financial crisis (0.83) then before (0.62), the difference is not statistically different. For all the others, even the magnitudes of the betas are very similar before and during the financial crisis. In summary, the results show that the curve betas of the sector ETFs do not increase during the market downward or bear market.

¹We also apply the models to different periods. The different sample selection does not influence the conclusions materially.

Comparing the betas during and after the financial crisis, the picture is quite different. For five XLB, XLE, XLK, XLU (utilities) and XLY (discretionary) out of the nine sector ETFs, the betas after the crisis are statistically significantly smaller than during the crisis at the 10% level. That is, the correlations between the asset and the market cumulative return curves generally decrease after the downward market.

Similarly, the results show that for five out of the nine sectors, the curve betas are significantly smaller after the financial crisis than before. For instance, the betas for XLB, XLK, and XLU are 0.87, 0.86 and 0.49 before the financial crisis. The corresponding betas are only 0.61, 0.47 and 0.28 during the “After” period. Interestingly, these are 30%, 45% and 43% smaller although the market conditions were similar during the two periods. However, betas for those sectors are dramatically different. It might indicate that investors’ behavior changes or may simply due to the portfolios holding different assets before and after the financial crisis. This issue might be partially revealed by testing individual stocks.

Table 2.3 reports the results for the 30 components of the Dow Jones Industrial Average Index. First, notice that all but one are statistically significantly different from zero with the estimate to standard error ratio ranging from 2.29 to 19.33 (the Beta of PG for the “Before” period has an estimate of 0.18 with standard error of 0.11). Second, 10 stocks have significantly different betas before and during the financial crisis. Among the 10 stocks, only two (AXP and BAC) have higher betas whereas the other eight have significantly lower betas during than before the financial crisis. This is consistent with the results for the ETFs. Third, one stock (MMM) has a significantly higher curve beta in the “After” period than the “During” period, 10 stocks have significantly lower curve betas in the “After” period than the “During” period whereas betas for the other 19 do not show statistically significant differences between the “After” and “During” periods. Fourth, a direct comparison of the “before” and “after” periods show that only one stock (GE) has a higher beta, 10 stocks have lower betas and betas of the other 19 are not significantly different between the two periods. In summary, the results for both the ETFs and the individual stocks indicate that the betas for most of those assets do not increase during the financial crisis. However, the betas of some of

the assets do decrease after the financial crisis. Therefore, the results are partially different from those reported in the literature for point-to-point returns.

Next we introduce the crude oil factor into the model and run a two factor model of the following form:

$$R_n(t) = \beta_0(t) + \beta_1 M_n(t) + \beta_2 C_n(t) + \varepsilon_n(t), \quad (2.24)$$

where, $C_n(t)$ is the CIDR of the Crude Oil Factor (CL).

Table 2.4 reports the results for the nine sector ETFs. In the “Before” period, the crude oil betas are only significantly different from zero for five ETFs. Except the beta for XLE (the Energy Sector ETF) which is 0.45, betas for all the other three ETFs are very small, even though they are statistically significantly different from zero (0.06, -0.04, -0.04, and -0.04 for XLB, XLP, XLV and XLY, respectively). For the “During” period, betas of the crude oil factor are significant for 6 sectors. The crude oil beta for XLE (0.44) is almost the same as for the “Before” period. The magnitudes of the other significant crude betas are much higher than before (0.21, 0.16, 0.07, 0.12 and 0.09 for the XLB, XLF, XLI, XLK and XLU). For the “After” period, all the crude oil betas are significant. The P-values of the pairwise joint tests show that jointly, 5, 7 and 9 ETFs have different betas for the three pair periods, respectively.

Comparing the results with Table 2.2, we observe that the additional significance is due to the significant differences of the crude oil factor. After introducing the crude oil factor, the number of significant P-Values for the “Before” and “During” periods increases from 1 to 5. This means that at least for four of the ETFs (XLB, XLF, XLP and XLU), the crude oil betas are different between the two periods. On the other hand, crude oil Betas for XLK are different (-0.03 with standard error 0.02 for the “Before” period and 0.12 with standard deviation of 0.04 for the “During” periods.) Since the test is a joint test and the test is significant in the one factor model setting, we can not directly conclude that the crude oil betas for XLE are different for the two periods. However, considering that the crude oil beta is not significant for the “Before” period and it is significant for the “During” period, we can safely conclude that the crude oil betas for the two periods are statistically different. Similarly, comparing the results for the other two paired periods, the results show that

adding the crude oil factor increases the number of significant results for all two paired periods. Interestingly, the crude oil betas for XLE are all significantly different from zero and consistent across all periods at roughly 0.44 – 0.45. The tests show they are not statistically different for all the paired periods. Broadly speaking, however, the results confirm that crude oil is an important factor for the sector ETFs and crude oil betas are time varying for most of the sectors.

Table 2.5 presents the results for the individual stocks. The results are consistent with the results for the ETFs. In general, observe that for most of the individual stocks except the three Energy stocks (AA, CVX, XOM), the crude oil betas are insignificant for the “Before” period. For the majority of the stocks, the crude oil betas are positive and significant for the “During” period. For the “After” period, the magnitudes of the crude oil betas increase further from the “During” period. The number of significant P-values of the test increased from 9, 11, 11 to 21, 13 and 29 and these findings support the conclusions obtained from the ETFs.

In summary, tests of the one factor and two factor models show that, for the ETFs and individual stocks, some have time-varying market betas and some have time-varying crude oil betas. However, in general, the results are not consistent with the findings in the literature that correlations between individual assets and the market increases during crisis periods. For most of the tested assets, market betas do not increase during the crisis period. We emphasize that the betas we study refer to intraday cumulative returns which can be impacted differently by a bear market than daily or monthly returns.

2.4.3 Asymmetric CIDR Betas for Bull and Bear Oil Price Periods

In this section, we consider only two samples corresponding to periods of increasing and decreasing oil prices. The two periods are based on crude oil price trends and are defined as: “Bull” (05/01/2007 – 06/30/2008) and “Bear” (07/01/2008 – 01/30/2009). These two periods are mostly occur during the financial crisis period and thus reduces the confounding of the two effects. The purpose of this section is to provide further insights on the effect of oil prices on cumulative intraday returns and demonstrate our methodology in a different context.

Table 2.6 exhibits the estimated results for the ETFs. Panel A presents the results for the one factor model of Equation (2.23). The P-Values of the test show that for all the ETFs, the market betas are not significantly different. When the oil factor is added, the P-values for six ETFs (XLF, XLI, XLK, XLP, XLU and XLV) are significantly different. For XLV, the crude oil beta for the first period is significant and negative, whereas for the second period it is insignificant. For the other five ETFs, $\hat{\beta}_{O2}$ is positive and statistically significantly different from zero whereas $\hat{\beta}_{O1}$ is statistically insignificant. For XLE, the crude oil factor is positive, stable and not different for the two periods ($\hat{\beta}_{O1} = 0.46$ and $\hat{\beta}_{O2} = 0.47$). For XLB, $\hat{\beta}_{O1} = 0.14$ and $\hat{\beta}_{O2} = 0.24$ and both of them are statistically different from zero, but are not statistically different from each other.

Table 2.7 shows analogous results for the individual stocks. Similarly, these results show that without the crude oil factor, only seven out of the 30 stocks have statistically asymmetric market beta during the two periods. When the crude oil factor is added, the number of significant P-Values increases to 23. For most of them, the crude oil price betas are negative and insignificant for the first period but positive and significant for the second period. The magnitudes of the oil factor also increase dramatically. For instance, for BAC, $\hat{\beta}_{O1} = -0.07$ and $\hat{\beta}_{O2} = 0.45$.

The results show that the crude oil beta for the majority of the assets is asymmetric for the two periods. For the “Bear” crude oil period, most of the crude oil betas are positive, significant with larger magnitudes than the “Bull” period. Since both of the “Bull” and “Bear” periods are within the downward stock market, the tests show that the crude oil factor is more important than the stock market. However, it can also be the case that during the second period, both the stock market and the crude oil price trended downward. Plausibly, the correlation between the crude oil and market factors are higher, leading to the higher crude oil betas.

2.4.4 Comparison to Point to Point Betas

Traditional financial models deal with point-to-point returns. Factor models for such returns are essentially scalar regression models. Our functional factor model reveals characteristics of intraday price curves which cannot be inferred using traditional factor models. It has a different information

content which focuses on the behavior of price curves within a trading day. The objective of this section is use a quantitative analysis to show that the information contained in point-to-point returns is often quite different from that obtained using the CIDR's and our functional approach. Therefore, in this section, betas obtained from point-to-point returns are estimated and compared with the CIDR betas. We consider the scalar model

$$R_n = \beta_0 + \sum_{j=1}^p \beta_j F_{nj} + \varepsilon_n, \quad (2.25)$$

where the intercept β_0 , factors F_{nj} 's and errors ε_n are all scalars. It can be verified that the entire test procedure described in Section 2.3 still holds for the scalar model (2.25), with obvious modifications. Two types of point-to-point returns are used. The first type are returns between the open and close of the market. These returns ignore the intraday evolution of the price curves and the overnight price change. The second type are the usual daily returns based on closing prices, i.e. returns between the close of trading day $n - 1$ and the close of trading day n . These returns take into account the overnight price change. To keep the length of the paper within limits, we report the results of the test for the ETFs only.

Table 2.8 reports the results for the one factor model. Comparing the results with Table 2.2, it is observed that the CIDR curve betas are smaller than the corresponding open-to-close betas which are smaller than the close-to-close return betas. For instance, for XLB, $\hat{\beta}_{M1}$ is 0.87, 0.91, and 1.23, for the curve, the open-to-close returns, and the close-to-close returns, respectively. This observation is consistent across all periods and for all the ETFs. Second, for some ETFs the pairwise test P-values for the estimated betas are different. For instance, for the "Before" and "During" periods, the test show that only one ETF has a statistically different CIDR beta ($\hat{\beta}_{M1} \neq \hat{\beta}_{M2}$), for the open-to-close return, three ETFs have different betas, and for the close-to-close beta, eight ETFs have significantly different betas. However, for the other two pair of periods, the differences are not that dramatic. For the "During" and "After" periods, the number of significant P-values are five for all three different return measures. Third, for some ETFs, the results with different returns

can be very different. For instance, the P-values of the Curve beta for XLI for all the three pairwise periods are insignificant. The close-to-close betas are statistically different for the “Before” and “During” as well as the “During” and “After” periods.

Table 2.9 reports the results for the two factor models for the three periods around the financial crisis. The results show similar patterns as the one factor model.

In general, the comparisons show that cumulative intraday returns and functional factor models reveal features different from point-to-point returns. Analysis based on functional regression models considers the relationships between the intraday prices of individual assets and factors. Consequently, these models are more relevant to investors taking positions within a trading day, i.e. day traders or high-frequency traders.

2.5 Finite sample performance

This section examines the test’s finite sample performance using sample sizes similar to those encountered in Section 2.4.2, and data generating processes that produce curves similar to the CIDR. The correct empirical sizes and sufficiently high power lend confidence to the findings reported in Section 2.4. The data generating processes are described and the empirical rejection rates are displayed.

The building block of artificial data is Brownian motion. Brownian motion is the standard model for price curves over short time intervals. In the simulations, Brownian motion is used to imitate the general shape of the CIDRs, as well as the error curves. We use B_n or B_m^* to denote the Brownian motions that are used to generate factors, and $\varepsilon_n, \varepsilon_m^*$ to denote the Brownian motions used to generate errors. Serial and cross-sectional dependence is generated by forming linear combinations of these processes. Specifically, we consider the case of $p = 2$, and set:

$$F_{n1}(t) = B_n^{(1)}(t) + 0.5B_{n-1}^{(1)}(t) + 0.25B_{n-2}^{(1)}(t), \quad (2.26)$$

$$F_{n2}(t) = B_n^{(2)}(t) + 0.5B_{n-1}^{(2)}(t) - 0.25B_{n-2}^{(1)}(t), \quad (2.27)$$

$$F_{m1}^*(t) = B_m^{*(1)}(t) + 0.5B_{m-1}^{*(1)}(t) + 0.25B_{m-2}^{*(1)}(t), \quad (2.28)$$

$$F_{m1}^*(t) = B_m^{*(2)}(t) + 0.5B_{m-1}^{*(2)}(t) - 0.25B_{m-2}^{*(1)}(t), \quad (2.29)$$

where the superscript “⁽¹⁾” and “⁽²⁾” on B_n and B_m represent different Brownian motions. The factors are thus both serially and cross-sectionally dependent. Letting $\beta_0(t) = 0$, which corresponds to the empirically observed value for most assets, we generate R_n and R_m^* using the following equations:

$$R_n(t) = \beta_1 F_{n1}(t) + \beta_2 F_{n2}(t) + \varepsilon_n(t), \quad (2.30)$$

$$R_m^*(t) = \beta_1^* F_{m1}^*(t) + \beta_2^* F_{m2}^*(t) + \varepsilon_m^*(t). \quad (2.31)$$

Under H_0 , we set $\beta_1 = \beta_1^*$ and $\beta_2 = \beta_2^*$. Under H_A , the data generation process for the second sample is modified by setting $\beta_1 \neq \beta_1^*$ and $\beta_2 \neq \beta_2^*$. To obtain relevant values of these parameters, three ETF's: XLB, XLK and XLU are selected to be used in model (2.24). Crude oil prices are used as the additional factor for the comparison between the during and after the crisis periods. For each of the three ETFs, we generate 328 market factor curves M_n^G , 328 crude oil factor curves C_n^G and 328 error curves $\hat{\varepsilon}_n^G$ such that they have approximately the same magnitude as the empirical curves M_n , C_n and $\hat{\varepsilon}_n$ for the period during the crisis, where $\hat{\varepsilon}_n$ is the residual function that can be calculated from the data. Analogously, 503 M_n^G , C_n^G and $\hat{\varepsilon}_n^G$ curves are generated for the period after the crisis. To assess the the empirical size of test, we generate the ETF curves as

$$R_n^G = \hat{\beta}_0(t) + \hat{\beta}_1^{ave} M_n^G(t) + \hat{\beta}_2^{ave} C_n^G(t) + \hat{\varepsilon}_n^G(t) \quad (2.32)$$

for each period, where $\hat{\beta}_1^{ave}$ and $\hat{\beta}_2^{ave}$ are the average of the coefficient estimates for the two periods.

To assess the empirical power, we use the data generation process

$$R_n^G = \hat{\beta}_0(t) + \hat{\beta}_1 M_n^G(t) + \hat{\beta}_2 C_n^G(t) + \hat{\varepsilon}_n^G(t) \quad (2.33)$$

for each period. We replicate this procedure $R = 3000$ times and obtain Monte Carlo empirical size and power. The simulation results exhibited in Table 2.11, demonstrate that the test exhibits excellent calibration size and has a very high power in practically relevant settings.

The empirical size evaluates the performance of a test when the null hypothesis is true, i.e. if there is no difference between the two samples. To illustrate, if we set the significance level at 5%, then for a perfect test we expect acceptance of the null in 95% of the simulation runs and rejection in 5% of them. This is never the case in practice. The actual percentage of rejections of the null is called the empirical size or the empirical type I error. Our test has a slight tendency to over reject, i.e. the probability of a type I error is generally slightly higher than the nominal level of the test.

Due to the slightly overinflated empirical size, when interpreting the P-values, one may expect that the true P-value may be slightly larger than the computed P-value. This distortion is, however, very small. For example, if a computed P-value is 0.045, the actual P-value can be slightly above 0.05. It must, however, be kept in mind that the traditional cut-off percentages, like 5% or 10%, are merely conventions; there is practically no difference between the statistical interpretation of a P-value of 4.5% or 5.5%. Comparing this test to other tests for complex data structures, the empirical size is remarkably close to the nominal size, and the computed P-values can be trusted.

The power of the test evaluates its performance if an alternative is true. In our case, power is evaluated for data simulated in such a way that the regression coefficients are different in the two samples. Power is the probability of rejecting the null hypothesis if an alternative is true. It depends on the alternative and on the significance level used to perform the test. If the difference between the samples is smaller, power will be smaller. If the significance level decreases, it becomes more and more difficult for the test to reject the null, so the power declines. We would like the power to be as close to 100% as possible. This is never the case in theory, because there is always a positive probability of accepting the null due to chance. If only a finite number of simulations are used to assess power, one can observe empirical power of 100% (the null is not accepted in any simulation

run). Our test has excellent power. If it is applied at the usual significance level of 5%, one can be almost certain to detect the difference between the two samples, if it exists.

2.6 Concluding Remarks

We propose functional factor models to study the relationship between the cumulative intraday return curves of individual stocks and factor curves that may drive price movement. To examine whether the betas change from period to period, we develop a statistical test. Based on general assumptions, we prove that the asymptotic distribution of the test statistic is a Chi-square distribution. We also prove theoretically that the power of the test approaches 100% asymptotically.

We apply the test to empirically investigate whether asymmetric betas exist for downward versus upward market conditions. We use the 30 stocks of the Dow Jones Industrial Average index and the Select Sector ETFs as the base assets. The tests on the one factor model show mixed results. For some assets betas are time varying and market betas of most assets do not increase during the financial crisis. However, most of the assets have lower market betas after the financial crisis.

When crude oil prices is added in the model, some assets exhibit asymmetric crude oil factor betas. Interestingly, the crude oil betas are not time varying for energy stocks and XLE, the energy ETF. This means that these assets are very strongly tied to the underlying commodity, and this connection is not influenced by the trends in the factor. We further examine the betas for the “Bull” and “Bear” crude oil price periods and the results also show that some assets have asymmetric crude oil factor betas.

A comparison of the results of the functional analysis model to scalar models using point-to-point returns show that the different return measures contain different information about the relationships between individual assets and the market factors. Considering the nature of the return curves, the models we propose and the results of our analysis are more useful for day traders and other higher frequency traders.

Table 2.1: List of Tickers

This table presents the tickers of the assets used in this study.

Ticker	Company/Sector	Ticker	Company/Sector
Panel A: Individual Stocks			
AA	Alcoa Inc.	AXP	American Express Company
BA	Boeing Co.	BAC	Bank of America Corp.
CAT	Caterpillar	CSCO	Cisco Systems Inc.
CVX	Chevron Corporation	DD	E.I. Du Pont De Nemours & Co.
DIS	Walt Disney Co.	GE	General Electric Company
HD	Home Depot Inc.	HPQ	Hewlett-Packard Company
IBM	International Business Machines Corp.	INTC	Intel Corporation
JNJ	Johnson & Johnson	JPM	JPMorgan Chase & Co.
KFT	Kraft Foods Inc.	KO	The Coca-Cola Co.
MCD	McDonald's Corporation	MMM	3M Co.
MRK	Merck & Co., Inc.	MSFT	Microsoft Corporation
PFE	Pfizer	PG	Procter & Gamble Co.
T	AT&T Inc.	TRV	Travelers Companies Inc.
UTX	United Technologies Co.	VZ	Verizon Communications Inc.
WMT	Wal-Mart Stores, Inc.	XOM	Exxon Mobil Corporation
Panel B: Sector ETFs			
XLY	Consumer Discretionary	XLP	Consumer Staples
XLE	Energy	XLF	Financials
XLV	Health Care	XLI	Industrials
XLB	Materials	XLK	Technology
XLU	Utilities		

Table 2.2: Asymmetric Beta around the Financial Crisis (One Factor) – ETFs

This table presents the estimation results for the model: $R_n(t) = \beta_0(t) + \beta_{Mi}M_n(t) + \varepsilon_n(t)$ for the sector ETFs. Here, i is the i^{th} periods. 1, 2, and 3 refer to the Before, During and After financial crisis periods, respectively. The periods are defined as: Before (06/13/2006 – 09/27/2007), During (11/01/2007 – 02/27/2009), and After (01/04/2010 – 12/30/2011) the financial crisis. SE_{Mi} refers to the standard error of the estimates. $P_{i,j}$ is the P-value of the statistical test with $H_0: \beta_{Mi} = \beta_{Mj}$. The superscripts ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{M3}$	SE_{M3}	$P_{1,2}$	$P_{2,3}$	$P_{1,3}$
XLB	0.87	0.07	0.76	0.05	0.61	0.04	0.21	0.02*	0.00***
XLE	0.78	0.09	0.81	0.07	0.65	0.04	0.78	0.04*	0.17
XLF	0.62	0.10	0.83	0.12	0.62	0.04	0.17	0.10	0.97
XLI	0.68	0.07	0.66	0.05	0.62	0.04	0.83	0.56	0.48
XLK	0.86	0.08	0.61	0.05	0.47	0.03	0.01**	0.02*	0.00***
XLP	0.37	0.05	0.30	0.06	0.23	0.03	0.35	0.23	0.01**
XLU	0.49	0.07	0.45	0.07	0.28	0.03	0.66	0.03*	0.00***
XLV	0.45	0.06	0.38	0.05	0.38	0.03	0.38	0.98	0.30
XLY	0.64	0.07	0.66	0.09	0.49	0.04	0.89	0.08*	0.05*

Table 2.3: Asymmetric Beta around the Financial Crisis (One Factor) – Stocks

This table presents the estimation results for the model: $R_n(t) = \beta_0(t) + \beta_{Mi}M_n(t) + \varepsilon_n(t)$ for the 30 Stocks. Here, i is the i^{th} periods. 1, 2, and 3 refer to the Before, During and After financial crisis periods, respectively. The periods are defined as: Before (06/13/2006 – 09/27/2007), During (11/01/2007 – 02/27/2009), and After (01/04/2010 – 12/30/2011) the financial crisis. $P_{i,j}$ is the P-value of the statistical test with $H_0: \beta_{Mi} = \beta_{Mj}$. The superscripts ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{M3}$	SE_{M3}	P _{1,2}	P _{2,3}	P _{1,3}
AA	0.94	0.12	0.83	0.14	0.71	0.06	0.55	0.42	0.09*
AXP	0.76	0.10	0.99	0.09	0.56	0.06	0.08*	0.00***	0.08*
BA	0.77	0.10	0.51	0.09	0.66	0.04	0.05*	0.10	0.33
BAC	0.41	0.09	0.90	0.21	0.56	0.09	0.03*	0.14	0.23
CAT	0.79	0.10	0.65	0.07	0.73	0.06	0.24	0.42	0.55
CSCO	0.68	0.10	0.58	0.07	0.52	0.04	0.45	0.41	0.14
CVX	0.46	0.09	0.62	0.05	0.46	0.04	0.11	0.01**	0.99
DD	0.69	0.15	0.59	0.06	0.58	0.05	0.55	0.94	0.51
DIS	0.61	0.12	0.76	0.06	0.56	0.04	0.29	0.01**	0.68
GE	0.49	0.05	0.64	0.14	0.62	0.05	0.29	0.89	0.05*
HD	0.68	0.09	0.62	0.09	0.37	0.06	0.64	0.01**	0.00***
HPQ	0.68	0.08	0.47	0.07	0.57	0.06	0.04*	0.25	0.26
IBM	0.39	0.08	0.49	0.06	0.38	0.03	0.31	0.10	0.94
INTC	0.74	0.08	0.51	0.08	0.41	0.04	0.04*	0.26	0.00***
JNJ	0.26	0.05	0.24	0.05	0.28	0.04	0.84	0.62	0.79
JPM	0.65	0.12	0.60	0.20	0.56	0.05	0.83	0.83	0.46
KFT	0.41	0.11	0.21	0.05	0.23	0.05	0.10	0.84	0.14
KO	0.23	0.07	0.29	0.06	0.28	0.04	0.43	0.82	0.50
MCD	0.52	0.11	0.35	0.05	0.19	0.03	0.17	0.00***	0.00***
MMM	0.56	0.09	0.36	0.06	0.58	0.03	0.06*	0.00***	0.84
MRK	0.59	0.07	0.42	0.06	0.41	0.04	0.08*	0.85	0.03*
MSFT	0.44	0.06	0.52	0.06	0.37	0.04	0.36	0.04*	0.34
PFE	0.60	0.08	0.35	0.06	0.36	0.04	0.01**	0.88	0.01**
PG	0.18	0.11	0.31	0.04	0.20	0.03	0.26	0.03*	0.87
T	0.47	0.07	0.61	0.06	0.29	0.03	0.13	0.00***	0.02*
TRV	0.49	0.14	0.32	0.14	0.30	0.04	0.37	0.90	0.18
UTX	0.72	0.09	0.54	0.06	0.56	0.04	0.08*	0.76	0.09*
VZ	0.48	0.11	0.56	0.06	0.22	0.03	0.53	0.00***	0.02*
WMT	0.54	0.10	0.31	0.07	0.17	0.04	0.07*	0.08*	0.00***
XOM	0.50	0.09	0.47	0.07	0.41	0.04	0.75	0.47	0.33

Table 2.4: Asymmetric Beta around the Financial Crisis (Two Factors) – ETFs

This table presents the estimation results for the model: $R_n(t) = \beta_0(t) + \beta_{Mi}M_n(t) + \beta_{Oi}CL_n(t) + \varepsilon_n(t)$ for the sector ETFs. Here, i is the i^{th} periods. 1, 2, and 3 refer to the Before, During and After financial crisis periods, respectively. The periods are defined as: Before (06/13/2006 – 09/27/2007), During (11/01/2007 – 02/27/2009), and After (01/04/2010 – 12/30/2011) the financial crisis. SE refers the standard error. $P_{i,j}$ is the P-value of the statistical test with $H_0: \beta_i = \beta_j$. The superscripts ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{O1}$	SE_{O1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{O2}$	SE_{O2}	$\hat{\beta}_{M3}$	SE_{M3}	$\hat{\beta}_{O3}$	SE_{O3}	P _{1,2}	P _{2,3}	P _{1,3}
XLB	0.86	0.07	0.06	0.03	0.68	0.05	0.21	0.04	0.47	0.04	0.34	0.04	0.01**	0.00***	0.00***
XLE	0.71	0.07	0.45	0.04	0.65	0.06	0.44	0.05	0.46	0.03	0.44	0.03	0.76	0.01**	0.00***
XLF	0.62	0.10	-0.01	0.02	0.77	0.12	0.16	0.07	0.51	0.04	0.27	0.04	0.01**	0.13	0.00***
XLI	0.68	0.07	0.01	0.02	0.63	0.05	0.07	0.03	0.54	0.04	0.21	0.04	0.15	0.01**	0.00***
XLK	0.86	0.08	-0.03	0.02	0.56	0.06	0.12	0.04	0.40	0.03	0.17	0.03	0.00***	0.02*	0.00***
XLP	0.38	0.05	-0.04	0.02	0.28	0.06	0.07	0.04	0.19	0.03	0.09	0.02	0.02*	0.42	0.00***
XLU	0.49	0.07	0.00	0.03	0.41	0.07	0.09	0.03	0.24	0.03	0.10	0.03	0.07*	0.07*	0.00***
XLV	0.46	0.06	-0.04	0.01	0.37	0.06	0.02	0.03	0.32	0.03	0.13	0.03	0.19	0.05*	0.00***
XLY	0.65	0.07	-0.04	0.02	0.64	0.09	0.06	0.06	0.40	0.04	0.20	0.04	0.19	0.05*	0.00***

Table 2.5: Asymmetric Beta around the Financial Crisis (Two Factors) -Stocks

This table presents the estimation results for the model: $R_n(t) = \beta_0(t) + \beta_{Mi}M_n(t) + \beta_{Oi}CL_n(t) + \varepsilon_n(t)$ for the sector ETFs. Here, i is the i^{th} periods. 1, 2, and 3 refer to the Before, During and After financial crisis periods, respectively. The periods are defined as: Before (06/13/2006 – 09/27/2007), During (11/01/2007 – 02/27/2009), and After (01/04/2010 – 12/30/2011) the financial crisis. SE refers the standard error. $P_{i,j}$ is the P-value of the statistical test with $H_0: \beta_i = \beta_j$. The superscripts ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{O1}$	SE_{O1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{O2}$	SE_{O2}	$\hat{\beta}_{M3}$	SE_{M3}	$\hat{\beta}_{O3}$	SE_{O3}	P _{1,2}	P _{2,3}	P _{1,3}
AA	0.93	0.12	0.18	0.05	0.71	0.12	0.43	0.07	0.53	0.06	0.48	0.06	0.01**	0.40	0.00***
AXP	0.76	0.10	0.01	0.03	0.94	0.09	0.16	0.07	0.46	0.06	0.26	0.05	0.04*	0.00***	0.00***
BA	0.77	0.10	0.03	0.04	0.47	0.09	0.14	0.06	0.57	0.05	0.25	0.04	0.07*	0.02*	0.00***
BAC	0.41	0.10	0.00	0.03	0.80	0.21	0.36	0.11	0.38	0.09	0.48	0.07	0.00***	0.17	0.00***
CAT	0.79	0.10	0.09	0.04	0.59	0.07	0.23	0.05	0.58	0.06	0.38	0.05	0.04*	0.06*	0.00***
CSCO	0.68	0.10	0.03	0.04	0.55	0.07	0.14	0.05	0.43	0.04	0.23	0.05	0.19	0.30	0.00***
CVX	0.44	0.08	0.40	0.04	0.53	0.05	0.33	0.04	0.33	0.03	0.33	0.03	0.29	0.00***	0.05*
DD	0.69	0.15	-0.01	0.03	0.55	0.06	0.15	0.04	0.48	0.05	0.29	0.04	0.01**	0.08*	0.00***
DIS	0.61	0.12	-0.04	0.04	0.72	0.06	0.12	0.04	0.52	0.04	0.11	0.04	0.01**	0.02*	0.01**
GE	0.49	0.05	-0.02	0.03	0.60	0.14	0.15	0.06	0.53	0.05	0.25	0.04	0.00***	0.38	0.00***
HD	0.67	0.09	0.01	0.04	0.60	0.09	0.06	0.06	0.28	0.06	0.22	0.05	0.69	0.01**	0.00***
HPQ	0.68	0.08	0.01	0.04	0.43	0.07	0.14	0.05	0.51	0.06	0.15	0.05	0.02*	0.57	0.04*
IBM	0.39	0.09	-0.01	0.03	0.45	0.06	0.13	0.04	0.33	0.04	0.14	0.03	0.00***	0.13	0.00***
INTC	0.74	0.08	-0.04	0.04	0.47	0.08	0.14	0.06	0.35	0.05	0.17	0.05	0.01**	0.37	0.00***
JNJ	0.26	0.05	-0.04	0.02	0.24	0.05	0.03	0.02	0.24	0.04	0.10	0.03	0.09*	0.10	0.00***
JPM	0.65	0.12	-0.02	0.03	0.56	0.21	0.17	0.09	0.42	0.05	0.36	0.05	0.14	0.19	0.00***
KFT	0.41	0.11	-0.01	0.04	0.19	0.05	0.07	0.03	0.19	0.06	0.11	0.03	0.05*	0.78	0.01**
KO	0.23	0.07	-0.05	0.03	0.29	0.06	0.02	0.03	0.25	0.04	0.08	0.03	0.16	0.38	0.00***
MCD	0.52	0.11	-0.01	0.03	0.34	0.05	0.03	0.03	0.14	0.03	0.12	0.03	0.28	0.00***	0.00***
MMM	0.56	0.09	0.01	0.04	0.34	0.06	0.06	0.04	0.51	0.03	0.19	0.04	0.10	0.00***	0.00***
MRK	0.59	0.07	-0.07	0.04	0.41	0.06	0.05	0.04	0.35	0.04	0.17	0.04	0.04*	0.12	0.00***
MSFT	0.44	0.06	0.02	0.03	0.49	0.06	0.08	0.04	0.31	0.04	0.16	0.05	0.31	0.05*	0.01**
PFE	0.60	0.08	-0.06	0.03	0.32	0.07	0.09	0.04	0.30	0.04	0.16	0.04	0.00***	0.41	0.00***
PG	0.18	0.11	-0.04	0.03	0.30	0.04	0.03	0.03	0.19	0.04	0.03	0.03	0.06*	0.08*	0.14
T	0.47	0.07	-0.03	0.04	0.57	0.06	0.17	0.05	0.25	0.03	0.09	0.03	0.00***	0.00***	0.00***
TRV	0.49	0.14	0.02	0.04	0.29	0.14	0.09	0.07	0.25	0.04	0.14	0.03	0.43	0.81	0.01**
UTX	0.72	0.09	0.01	0.04	0.51	0.06	0.10	0.04	0.51	0.04	0.15	0.04	0.04*	0.71	0.01**
VZ	0.48	0.11	-0.01	0.04	0.52	0.07	0.12	0.04	0.18	0.04	0.10	0.03	0.07*	0.00***	0.01**
WMT	0.54	0.10	-0.08	0.03	0.31	0.07	0.01	0.04	0.16	0.05	0.04	0.03	0.04*	0.19	0.00***
XOM	0.49	0.08	0.38	0.03	0.39	0.07	0.28	0.04	0.31	0.04	0.27	0.03	0.04*	0.46	0.00***

Table 2.6: Bull vs. Bear Oil Price Periods: ETFs

This table presents the estimation results for the ETFs. The periods are defined as: Bull oil price (05/01/2007 – 06/30/2008), and Bear oil price (07/01/2008 – 01/30/2009). The number of sample days for the two 294 and 145 days, respectively. $P_{1,2}$ is the P-value of the statistical test with $H_0: \beta_1 = \beta_2$. ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{O1}$	SE_{O1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{O2}$	SE_{O2}	$P_{1,2}$
Panel A: Single Factor Model									
XLB	0.73	0.06			0.77	0.06			0.64
XLE	0.67	0.06			0.82	0.08			0.15
XLF	0.80	0.10			0.77	0.15			0.86
XLI	0.60	0.07			0.65	0.06			0.55
XLK	0.68	0.07			0.59	0.07			0.35
XLP	0.26	0.04			0.30	0.08			0.67
XLU	0.31	0.07			0.49	0.09			0.11
XLV	0.38	0.06			0.37	0.07			0.93
XLY	0.65	0.07			0.64	0.11			0.96
Panel B: Two Factor Model									
XLB	0.73	0.07	0.14	0.04	0.68	0.06	0.24	0.06	0.43
XLE	0.66	0.05	0.46	0.04	0.64	0.08	0.47	0.07	0.95
XLF	0.80	0.10	-0.10	0.06	0.67	0.16	0.24	0.10	0.01**
XLI	0.60	0.07	-0.02	0.03	0.61	0.07	0.10	0.04	0.01**
XLK	0.68	0.07	-0.03	0.04	0.52	0.07	0.18	0.06	0.01**
XLP	0.26	0.04	-0.04	0.04	0.26	0.08	0.10	0.05	0.03*
XLU	0.31	0.07	-0.01	0.04	0.45	0.09	0.11	0.04	0.06*
XLV	0.38	0.06	-0.09	0.03	0.34	0.07	0.07	0.05	0.01**
XLY	0.65	0.07	-0.07	0.04	0.60	0.12	0.10	0.08	0.11

Table 2.7: Bull vs. Bear Oil Price Periods: Individual Stocks

This table presents the estimation results for the individual stocks. The periods are defined as: Bull oil price (05/01/2007 – 06/30/2008), and Bear oil price (07/01/2008 – 01/30/2009). The number of sample days for the two periods are 294 and 145 days, respectively. $P_{1,2}$ is the P-value of the statistical test with $H_0: \beta_1 = \beta_2$. ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively. To save space, only the significant results are reported.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{O1}$	SE_{O1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{O2}$	SE_{O2}	$P_{1,2}$
Panel A: Single Factor Model									
CVX	0.50	0.06			0.65	0.07			0.09*
DIS	0.50	0.07			0.83	0.08			0.00***
HD	0.80	0.13			0.51	0.10			0.08*
INTC	0.72	0.07			0.43	0.10			0.01**
MMM	0.53	0.07			0.26	0.07			0.01**
PFE	0.48	0.06			0.30	0.08			0.07*
PG	0.15	0.06			0.34	0.05			0.02*
Panel B: Two Factor Model									
AA	0.93	0.08	0.23	0.07	0.57	0.15	0.46	0.09	0.03*
AXP	1.07	0.12	-0.08	0.06	0.81	0.12	0.26	0.10	0.01**
BAC	0.63	0.14	-0.07	0.07	0.61	0.24	0.45	0.15	0.00***
CAT	0.66	0.08	0.07	0.05	0.50	0.10	0.25	0.06	0.05*
CSCO	0.58	0.08	-0.04	0.06	0.48	0.09	0.21	0.07	0.01**
DIS	0.50	0.07	-0.03	0.04	0.78	0.08	0.19	0.06	0.00***
GE	0.63	0.05	-0.08	0.04	0.51	0.18	0.24	0.09	0.00***
HD	0.80	0.13	-0.08	0.06	0.47	0.11	0.13	0.07	0.04*
HPQ	0.60	0.07	-0.00	0.06	0.38	0.09	0.19	0.07	0.03*
IBM	0.49	0.08	-0.03	0.04	0.39	0.07	0.17	0.05	0.01**
INTC	0.72	0.07	-0.05	0.05	0.37	0.10	0.20	0.07	0.00***
JNJ	0.19	0.04	-0.07	0.03	0.24	0.07	0.05	0.03	0.01**
JPM	0.72	0.13	-0.13	0.07	0.41	0.28	0.27	0.13	0.02*
KFT	0.34	0.07	-0.01	0.04	0.15	0.06	0.12	0.04	0.03*
KO	0.23	0.06	-0.10	0.04	0.29	0.07	0.05	0.05	0.04*
MMM	0.53	0.07	-0.05	0.03	0.23	0.07	0.11	0.05	0.00***
MRK	0.38	0.07	-0.07	0.05	0.41	0.08	0.13	0.06	0.03*
MSFT	0.47	0.06	-0.04	0.04	0.50	0.09	0.14	0.06	0.01**
PFE	0.48	0.05	-0.10	0.03	0.25	0.08	0.17	0.05	0.00***
PG	0.14	0.06	-0.07	0.03	0.32	0.05	0.04	0.04	0.00***
T	0.58	0.07	-0.05	0.04	0.54	0.08	0.23	0.07	0.00***
VZ	0.54	0.09	-0.05	0.05	0.48	0.08	0.17	0.06	0.01**
WMT	0.35	0.08	-0.10	0.04	0.24	0.09	0.04	0.05	0.08*

Table 2.8: Two Sample Test Results – ETFs -Point to Point

This table presents the estimation results for the point to point return series of the ETFs. The periods are defined as: Before (06/13/2006 – 09/27/2007), During (11/01/2007 – 02/27/2009), and After (01/04/2010 – 12/30/2011) the financial crisis. The number of sample days for the three periods are 326, 328 and 503 days, respectively. $P_{i,j}$ is the P-value of the statistical test with $H_0: \beta_i = \beta_j$. ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{M3}$	SE_{M3}	$P_{1,2}$	$P_{2,3}$	$P_{1,3}$
Panel A: Open to Close Return									
XLB	0.91	0.07	0.88	0.03	0.83	0.04	0.66	0.39	0.34
XLE	0.94	0.07	0.99	0.06	0.82	0.04	0.62	0.02*	0.13
XLF	0.86	0.07	1.19	0.10	0.92	0.04	0.01**	0.01**	0.53
XLI	0.79	0.05	0.76	0.03	0.76	0.03	0.56	1.00	0.57
XLK	0.89	0.06	0.74	0.03	0.60	0.03	0.02*	0.00***	0.00***
XLP	0.50	0.03	0.43	0.04	0.35	0.02	0.19	0.09*	0.00***
XLU	0.74	0.07	0.57	0.04	0.41	0.03	0.04*	0.00***	0.00***
XLV	0.57	0.04	0.54	0.04	0.50	0.03	0.51	0.46	0.13
XLY	0.79	0.05	0.83	0.07	0.67	0.04	0.56	0.02*	0.04*
Panel B: Close to Close Return									
XLB	1.23	0.05	1.08	0.03	1.25	0.02	0.01**	0.00***	0.76
XLE	1.11	0.06	1.27	0.07	1.25	0.03	0.08*	0.78	0.04*
XLF	1.22	0.04	1.57	0.10	1.35	0.03	0.00***	0.03*	0.02*
XLI	1.02	0.03	0.92	0.02	1.15	0.01	0.01**	0.00***	0.00***
XLK	1.02	0.04	0.93	0.02	0.94	0.01	0.04*	0.59	0.07*
XLP	0.68	0.02	0.57	0.03	0.57	0.01	0.00***	0.96	0.00***
XLU	0.91	0.06	0.72	0.05	0.64	0.02	0.02*	0.08*	0.00***
XLV	0.73	0.03	0.66	0.03	0.76	0.02	0.06*	0.00***	0.42
XLY	1.05	0.03	1.01	0.03	1.02	0.02	0.29	0.77	0.32

Table 2.9: Two sample test for the comparison two Factor Models -ETFs- Point to Point

This table presents the estimation results for the point to point return series of the ETFs. The periods are defined as: Before (06/13/2006 – 09/27/2007), During (11/01/2007 – 02/27/2009), and After (01/04/2010 – 12/30/2011) the financial crisis. The number of sample days for the three periods are 326, 328 and 503 days, respectively. $P_{i,j}$ is the P-value of the statistical test with $H_0: \beta_i = \beta_j$. ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{O1}$	SE_{O1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{O2}$	SE_{O2}	$\hat{\beta}_{M3}$	SE_{M3}	$\hat{\beta}_{O3}$	SE_{O3}	$P_{1,2}$	$P_{2,3}$	$P_{1,3}$
Panel A: Open to Close Return															
XLB	0.90	0.07	0.05	0.03	0.83	0.03	0.15	0.04	0.70	0.04	0.27	0.04	0.10	0.02*	0.00***
XLE	0.87	0.07	0.39	0.04	0.88	0.06	0.33	0.05	0.65	0.03	0.34	0.03	0.67	0.00***	0.00***
XLF	0.87	0.07	-0.03	0.02	1.16	0.10	0.10	0.06	0.83	0.05	0.17	0.04	0.00***	0.02*	0.00***
XLI	0.79	0.05	0.00	0.02	0.75	0.03	0.02	0.03	0.69	0.04	0.13	0.03	0.75	0.02*	0.00***
XLK	0.90	0.06	-0.04	0.02	0.72	0.03	0.06	0.03	0.55	0.03	0.10	0.03	0.00***	0.00***	0.00***
XLP	0.51	0.03	-0.04	0.02	0.43	0.04	0.02	0.03	0.33	0.03	0.04	0.02	0.14	0.15	0.00***
XLU	0.74	0.07	0.00	0.03	0.56	0.04	0.04	0.03	0.38	0.03	0.07	0.03	0.06*	0.00***	0.00***
XLV	0.58	0.04	-0.04	0.01	0.55	0.04	-0.03	0.03	0.46	0.03	0.08	0.03	0.79	0.01**	0.00***
XLY	0.80	0.05	-0.05	0.02	0.83	0.07	0.00	0.05	0.60	0.04	0.13	0.04	0.34	0.02*	0.00***
Panel B: Close to Close Return															
XLB	1.22	0.06	0.05	0.02	1.01	0.03	0.12	0.03	1.19	0.03	0.07	0.02	0.00***	0.00***	0.86
XLE	1.04	0.06	0.40	0.03	1.08	0.07	0.35	0.04	1.06	0.03	0.20	0.02	0.58	0.00***	0.00***
XLF	1.23	0.05	-0.06	0.02	1.65	0.10	-0.15	0.04	1.40	0.04	-0.05	0.02	0.00***	0.01**	0.00***
XLI	1.02	0.03	-0.02	0.01	0.93	0.02	-0.02	0.02	1.16	0.02	-0.01	0.01	0.05*	0.00***	0.00***
XLK	1.03	0.04	-0.06	0.01	0.93	0.02	0.00	0.01	0.97	0.02	-0.03	0.01	0.01**	0.12	0.35
XLP	0.68	0.02	-0.04	0.01	0.60	0.03	-0.05	0.01	0.59	0.02	-0.02	0.01	0.03*	0.31	0.00***
XLU	0.91	0.07	0.00	0.02	0.72	0.05	0.01	0.03	0.62	0.02	0.01	0.01	0.06	0.17	0.00***
XLV	0.74	0.03	-0.05	0.01	0.69	0.03	-0.06	0.02	0.78	0.02	-0.02	0.01	0.31	0.00***	0.03*
XLY	1.06	0.03	-0.05	0.01	1.04	0.03	-0.06	0.02	1.06	0.02	-0.04	0.01	0.77	0.61	0.95

Table 2.10: Bull vs. Bear Oil Price Periods: ETFs -Point to Point

This table presents the estimation results for the point to point return series of the ETFs. The periods are defined as: Bull oil price (05/01/2007 – 06/27/2008), and Bear oil price (07/01/2008 – 01/30/2009). The number of sample days for the two 294 and 145 days, respectively. $P_{1,2}$ is the P-value of the statistical test with $H_0: \beta_1 = \beta_2$. ***, **, *, and * indicate significant at 0.1%, 1%, 5% and 10%, respectively. To save space, only the significant results are reported.

Symbol	$\hat{\beta}_{M1}$	SE_{M1}	$\hat{\beta}_{O1}$	SE_{O1}	$\hat{\beta}_{M2}$	SE_{M2}	$\hat{\beta}_{O2}$	SE_{O2}	$P_{1,2}$
Panel A: Single Factor Model–OC									
No Significant Estimate									
Panel B: One Factor Model–CC									
XLB	1.17	0.05			1.07	0.03			0.06*
XLE	0.99	0.05			1.34	0.07			0.00***
XLF	1.50	0.04			1.54	0.12			0.79
XLI	0.99	0.02			0.90	0.02			0.01**
XLK	1.00	0.03			0.91	0.03			0.03*
Panel C: Two Factor Model–OC									
XLF	1.12	0.07	-0.10	0.05	1.10	0.12	0.16	0.09	0.03*
XLK	0.80	0.04	-0.03	0.03	0.69	0.04	0.08	0.04	0.05*
Panel D: Two Factor Model–CC									
XLB	1.17	0.05	0.19	0.03	0.99	0.04	0.12	0.03	0.00***
XLI	0.99	0.02	-0.01	0.01	0.92	0.03	-0.03	0.03	0.05*
XLK	1.00	0.03	-0.03	0.02	0.90	0.03	0.01	0.01	0.04*
XLV	0.60	0.03	-0.06	0.02	0.71	0.04	-0.06	0.02	0.09*
XLY	1.03	0.03	-0.11	0.02	1.03	0.04	-0.04	0.03	0.06*

Table 2.11: Monte Carlo Simulation Results of the test statistics $T_{M,N}$

This table presents the Monte Carlo Simulation Results using three ETFs as examples.

Sig. Level	XLB	XLK	XLU
Panel A: Empirical Size			
10%	11.3%	9.8%	11.3%
5%	5.9%	5.6%	5.4%
2.5%	2.8%	2.7%	2.6%
1%	1.0%	1.2%	1.2%
Panel B: Empirical Power			
10%	100.0%	100%	97.4%
5%	100.0%	99.8%	94.4%
2.5%	100.0%	99.7%	90.8%
1%	100.0%	99.2%	85.3%

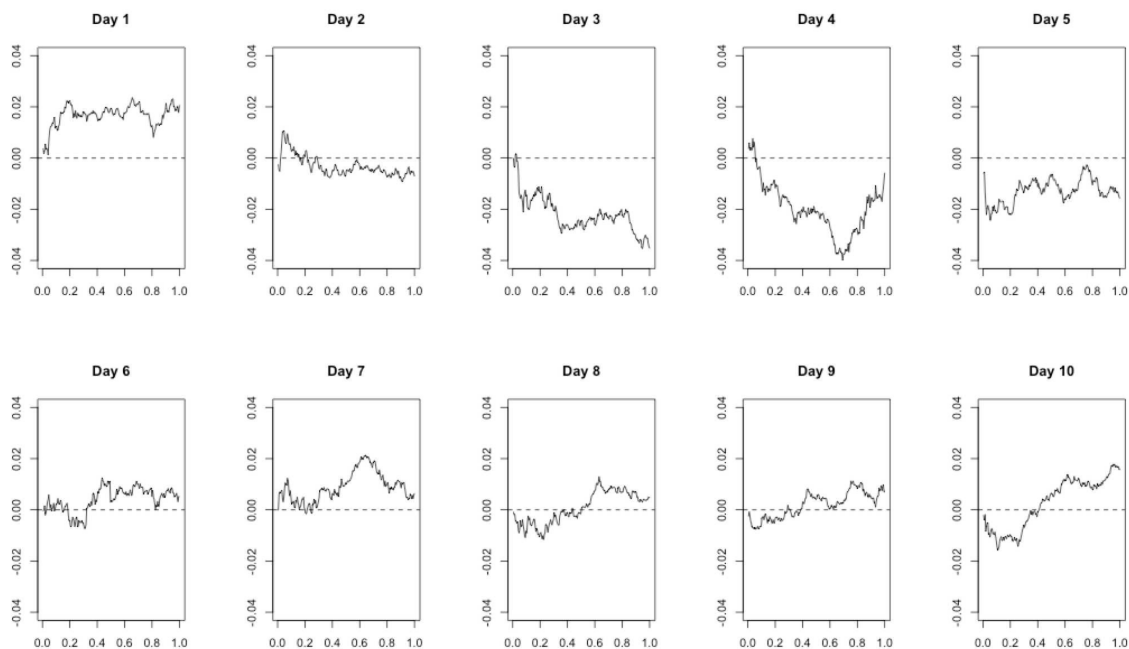


Figure 2.1: Ten Consecutive CIDRs on JP Morgan Chase Stock. This figure shows ten consecutive days CIDR on the JP Morgan Chase Stock. While each curve is a realization of a non-stationary stochastic process, the sequence of curves can be assumed to be a realization of a stationary time series. The curves have approximately the same distribution as the Brownian motion.

2.7 Proofs

2.7.1 Proof of Theorem 5

Proof. Since we assume that the two samples are independent, the limits in (2.20) and (2.21) are independent. Combining (2.20), (2.21) and (2.18), we obtain

$$\begin{aligned}\sqrt{N+M}(\hat{\beta} - \hat{\beta}^*) &= \sqrt{\frac{N+M}{N}}\sqrt{N}(\hat{\beta} - \beta) - \sqrt{\frac{N+M}{M}}\sqrt{M}(\hat{\beta}^* - \beta) \\ &\xrightarrow{d} \frac{1}{\sqrt{\theta}}\mathbf{F}^{-1}\mathbf{W} + \frac{1}{\sqrt{1-\theta}}(\mathbf{F}^*)^{-1}\mathbf{W}^*.\end{aligned}\quad (2.34)$$

Define

$$\hat{\mathbf{D}}_{N,M} = \frac{N+M}{N}\hat{\mathbf{F}}^{-1}\hat{\mathbf{\Gamma}}\hat{\mathbf{F}}^{-1} + \frac{N+M}{M}(\hat{\mathbf{F}}^*)^{-1}\hat{\mathbf{\Gamma}}^*(\hat{\mathbf{F}}^*)^{-1}$$

It was shown in [2], Lemma 4, that $\hat{\mathbf{F}} \xrightarrow{a.s.} \mathbf{F}$. By assumption, $\hat{\mathbf{\Gamma}} \xrightarrow{P} \mathbf{\Gamma}$. With analogous relations for the second sample, we thus obtain

$$\hat{\mathbf{D}}_{N,M} \xrightarrow{P} \frac{1}{\theta}\mathbf{F}^{-1}\mathbf{\Gamma}\mathbf{F}^{-1} + \frac{1}{1-\theta}(\mathbf{F}^*)^{-1}\mathbf{\Gamma}^*(\mathbf{F}^*)^{-1}.\quad (2.35)$$

Observe that \mathbf{W} and \mathbf{W}^* are two p -dimensional random vectors satisfying

$$\begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \sim \mathbf{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}^* \end{bmatrix}\right).\quad (2.36)$$

Since the matrices \mathbf{F} and \mathbf{F}^* defined by (2.19) are nonsingular, we can conclude that

$$\frac{1}{\sqrt{\theta}}\mathbf{F}^{-1}\mathbf{W} + \frac{1}{\sqrt{1-\theta}}(\mathbf{F}^*)^{-1}\mathbf{W}^* \sim N\left(\mathbf{0}, \frac{1}{\theta}\mathbf{F}^{-1}\mathbf{\Gamma}\mathbf{F}^{-1} + \frac{1}{1-\theta}(\mathbf{F}^*)^{-1}\mathbf{\Gamma}^*(\mathbf{F}^*)^{-1}\right),\quad (2.37)$$

where the variance matrix of $\frac{1}{\sqrt{\theta}}\mathbf{F}^{-1}\mathbf{W} + \frac{1}{\sqrt{1-\theta}}(\mathbf{F}^*)^{-1}\mathbf{W}^*$ has rank p , also $\mathbf{\Gamma}$ and $\mathbf{\Gamma}^*$ are assumed to be full rank.

Now Let

$$\mathbf{Y}_p := \frac{1}{\sqrt{\theta}}\mathbf{F}^{-1}\mathbf{W} + \frac{1}{\sqrt{1-\theta}}(\mathbf{F}^*)^{-1}\mathbf{W}^*$$

and

$$\Sigma_p := \frac{1}{\theta} \mathbf{F}^{-1} \mathbf{\Gamma} \mathbf{F}^{-1} + \frac{1}{1-\theta} (\mathbf{F}^*)^{-1} \mathbf{\Gamma}^* (\mathbf{F}^*)^{-1},$$

where Σ_p is nonsingular. Then (2.37) becomes $\mathbf{Y}_p \sim N(\mathbf{0}, \Sigma_p)$. This follows from the identity $\mathbf{Y}_p^T \Sigma_p^{-1} \mathbf{Y}_p \sim \chi^2(p)$. By Slutsky's Theorem, the claim of the theorem follows by combining (2.34) and (2.35). \square

2.7.2 Proof of Theorem 6

Proof. Let

$$\begin{aligned} \mathbf{B}_{N,M} &:= \sqrt{\frac{N+M}{N}} \sqrt{N} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) - \sqrt{\frac{N+M}{M}} \sqrt{M} (\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^*), \\ \mathbf{B} &:= \frac{1}{\sqrt{\theta}} \mathbf{F}^{-1} \mathbf{W} + \frac{1}{\sqrt{1-\theta}} (\mathbf{F}^*)^{-1} \mathbf{W}^*, \\ \mathbf{D} &:= \frac{1}{\theta} \mathbf{F}^{-1} \mathbf{\Gamma} \mathbf{F}^{-1} + \frac{1}{1-\theta} (\mathbf{F}^*)^{-1} \mathbf{\Gamma}^* (\mathbf{F}^*)^{-1}, \end{aligned}$$

then by (2.34) and (2.37), $\mathbf{B}_{N,M} \xrightarrow{d} \mathbf{B}$, where $\mathbf{B} \sim N(\mathbf{0}, \mathbf{D})$. Since $\boldsymbol{\beta}^* = \boldsymbol{\beta} + \boldsymbol{\eta}$, we have

$$\begin{aligned} \sqrt{N+M} (\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^*) &= \sqrt{\frac{N+M}{N}} \sqrt{N} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) - \sqrt{\frac{N+M}{M}} \sqrt{M} (\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^* + \boldsymbol{\eta}) \\ &= \sqrt{\frac{N+M}{N}} \sqrt{N} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) - \sqrt{\frac{N+M}{M}} \sqrt{M} (\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^*) - \sqrt{N+M} \boldsymbol{\eta} \\ &= \mathbf{B}_{N,M} - \sqrt{N+M} \boldsymbol{\eta}. \end{aligned} \tag{2.38}$$

Let $\tilde{T}_{N,M} := \mathbf{B}_{N,M}^T \hat{\mathbf{D}}_{N,M}^{-1} \mathbf{B}_{N,M}$, then by Theorem 5, $\tilde{T}_{N,M} \xrightarrow{d} \chi^2(p)$. Now we have

$$\begin{aligned} T_{N,M} &= (N+M) (\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^*)^T (\hat{\mathbf{D}}_{M,N})^{-1} (\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^*) \\ &= (\mathbf{B}_{N,M} - \sqrt{N+M} \boldsymbol{\eta})^T (\hat{\mathbf{D}}_{M,N})^{-1} (\mathbf{B}_{N,M} - \sqrt{N+M} \boldsymbol{\eta}) \\ &= \mathbf{B}_{N,M}^T (\hat{\mathbf{D}}_{M,N})^{-1} \mathbf{B}_{N,M} - 2\sqrt{N+M} \boldsymbol{\eta}^T (\hat{\mathbf{D}}_{M,N})^{-1} \mathbf{B}_{N,M} \\ &\quad + (N+M) \boldsymbol{\eta}^T (\hat{\mathbf{D}}_{M,N})^{-1} \boldsymbol{\eta} \\ &= \tilde{T}_{N,M} - 2\sqrt{N+M} \boldsymbol{\eta}^T (\hat{\mathbf{D}}_{M,N})^{-1} \mathbf{B}_{N,M} + (N+M) \boldsymbol{\eta}^T (\hat{\mathbf{D}}_{M,N})^{-1} \boldsymbol{\eta}. \end{aligned} \tag{2.39}$$

It follows that

$$\begin{aligned}
(N+M)^{-1}T_{N,M} &= (N+M)^{-1}\tilde{T}_{N,M} - 2(\sqrt{N+M})^{-1}\boldsymbol{\eta}^T(\widehat{\boldsymbol{D}}_{M,N})^{-1}\mathbf{B}_{N,M} \\
&\quad + \boldsymbol{\eta}^T(\widehat{\boldsymbol{D}}_{M,N})^{-1}\boldsymbol{\eta} \\
&\stackrel{d}{\rightarrow} (N+M)^{-1}\chi^2(p) - 2(\sqrt{N+M})^{-1}\boldsymbol{\eta}^T\boldsymbol{D}^{-1}\mathbf{B} + \boldsymbol{\eta}^T\boldsymbol{D}^{-1}\boldsymbol{\eta} \\
&\stackrel{d}{\rightarrow} \boldsymbol{\eta}^T\boldsymbol{D}^{-1}\boldsymbol{\eta}, \tag{2.40}
\end{aligned}$$

where $\boldsymbol{\eta}^T\boldsymbol{D}^{-1}\mathbf{B} \sim N(0, \boldsymbol{\eta}^T\boldsymbol{D}^{-1}\boldsymbol{\eta})$. This implies $T_{N,M} \xrightarrow{P} \infty$. □

Chapter 3

Risk Analysis of Cumulative Intraday Return Curves

3.1 Introduction

A fundamental concept of quantitative finance is the point-to-point return. For example, daily returns are defined as

$$r_n = 100(\ln P_n - \ln P_{n-1}) \approx 100 \frac{P_n - P_{n-1}}{P_{n-1}}, \quad (3.1)$$

where P_n is the closing price on trading day n . The unit of time can be day, month, or year, with the classical finance theory initiated by [33], [34], [35] and [36] pertaining to annual returns. This Chapter is concerned with daily curves of the cumulative intraday returns defined as follows.

Definition 7. Suppose $P_n(t)$, $n = 1, \dots, N$, is the price of a financial asset at time t on trading day n . The functions

$$R_n(t) = 100[\ln P_n(t) - \ln P_n(t_0)], \quad n = 1, \dots, N,$$

are called the cumulative intraday returns (CIDRs).

In practice, the CIDRs can be computed only at discrete time points, but the resolution can be very fine for frequently traded assets and indexes. In this research, we use one-minute resolution with t_0 corresponding to the first minute after the market opening; $P_n(t)$ is the average of the maximum and minimum price within the minute. For illustration, five consecutive CIDR curves are displayed in Figure 3.1. These curves show how the return accumulates throughout a trading day. Since $R_n(t) \approx 100(P_n(t) - P_n(t_0))/P_n(t_0)$, with $P_n(t_0)$ being a constant for a given day n , the price and the CIDR curves look very similar. However, the CIDR curves are scale and level normalized, so they form a stationary sequence in a function space, [30]. We view CIDR curves

as a time series of functions, and focus is on the analysis of their shapes. For this reason, we adopt the general framework of functional data analysis. While a substantial body of work on various aspects of time series of functions has accumulated, as we discuss below, we are not aware of any work on the *extremes* of such time series. This paper is an attempt to provide a possible definition of extreme curves and study their properties in the context of CIDRs. Extreme value analysis of point-to-point returns occupies a prominent place in econometric research due to its direct relation to risk analysis, but no work has been done on the extreme shapes of the CIDR *curves*.

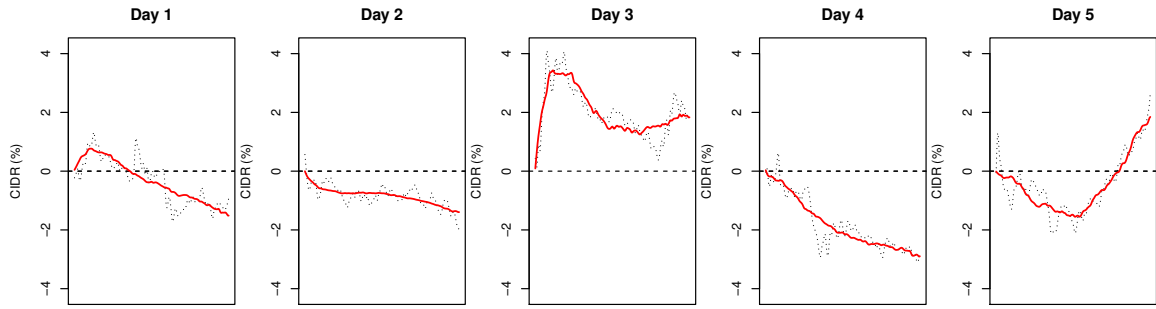


Figure 3.1: Five Consecutive CIDRs (sector ETF XLF). The noisy curves are the raw CIDRs, the superimposed thick curves are the corresponding CIDRs reconstructed from the first three FPCs. Details are explained in Section 3.2.

The CIDR curves describe how the return on an investment evolves with time over a relatively short period. The information contained in them is different than the information provided by minute-by-minute or trade-by-trade returns, and may be more relevant in situations where positions are held for a period of few hours within a trading day. Risk measures computed for point-to-point returns may be not sufficient for evaluating risks for portfolios held by such trades. For example, valuations of certain options are impacted by the shape the price curve may take over some interval, and knowing the probabilities of pre-specified unusual shapes may be useful. The type of risk we study is different from risk related to micro-structure noise. The latter is relevant to high frequency trades executed within seconds. The risk inherent in the shapes of cumulative returns is relevant to positions held at least for several hours. Micro-structure noise is relevant to professional traders, whereas smooth within day price evolution is relevant to a broader, and

growing, group of investors owning, for example, Exchange Traded Funds, which can be bought or sold within a trading day. This time horizon is in turn shorter than that of longer term investors, holding, for example, positions in Mutual Funds.

Our objective is to develop a possible general framework, we do not develop any asymptotic theory at this stage, but rather focus on illustrating the concepts by means of various graphical displays.

The paper is organized as follows. In the remainder of the introduction, we cite a few selected monographs and papers related to the themes of our work: functional data analysis and risk analysis of returns. These fields of research are large, so our small review is necessarily selective and focused on work most relevant to our research. Section 3.2 presents statistical preliminaries, including the most relevant methodology of functional data analysis and extreme value theory. Using these tools, we conduct a risk analysis of several aspects of CIDRs in Section 3.3. We focus on the impact of the financial crisis of 2008 on the extremal behavior of the CIDR curves. Large tables and graphical displays are collected in Section 3.4, so that they do not interrupt the narrative.

General introductions to functional data analysis (FDA) are given in [37] and [38]. Methodology based on functional principal components, which we use extensively here, is the focus of [31]. Early results on linear functional time series, including their prediction are presented in [39]. Ideas of FDA have been applied to inference for CIDRs by [40], [41], [2], [42], [43] and [44].

Extreme value theory (EVT) with a view toward risk analysis is presented in [45] and [46], monographs which define and study the value-at-risk (VaR) and the expected shortfall (ES) (we review these concepts in Section 3.2). Extreme behavior of *point-to-point* returns has been studied in many papers. For example, [47] applied univariate EVT techniques to a series of daily exchange rates of Canadian/U.S. dollars over a 5-year period (1995-2000), [48] compared the performance of the extreme value theory in VaR calculations to other well-known modeling techniques and concluded that the EVT provides an approach which is robust and easy to implement. [4] used the extreme value theory to compute VaR and ES and related confidence intervals of several major stock market indices for the period 1960–2004. [49] investigated VaR and ES for several market

indexes during the 2008 financial crisis. Other papers discussing the tail behavior of financial series include [50], [51], [52], [53], [54] and [55]. [56] provides a critical review of the theoretical underpinnings of extreme value theory and a survey of some major applications of extreme value theory to finance. Intraday VaR was initially discussed by [57] who estimates a conditional parametric VaR in which intraday volatility is modeled using different specifications of GARCH, and RiskMetrics models on 15- and 30-minute returns of three stocks traded on the New York Stock Exchange. More recently, [58] proposed an approach to estimate intraday Value at Risk (IVaR) using tick-by-tick data. Value-at-risk or Conditional Value-at-risk using intraday price data contains richer information than the traditional daily data based measures and are proper for shorter term portfolios. Our analysis is different, it focuses on extreme shapes of the intraday return curves.

3.2 Statistical Preliminaries

This section introduces the requisite statistical background in FDA and EVT. Readers familiar with these fields may scan this section to become acquainted with the notation we use.

3.2.1 Background on Functional Data Analysis

Square integrable functions We assume that all functions are defined on a compact interval, which, for simplicity, we take to be the unit interval $[0, 1]$. A function f is said to be square integrable if $\int_0^1 f^2(t)dt < \infty$. We denote the set of all the square integrable functions by L^2 . It is a Hilbert space with the inner product

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt, \quad f, g \in L^2.$$

We say that functions f and g are orthogonal if $\langle f, g \rangle = 0$. The inner product also allows us to define the *norm* as

$$\|f\| = \sqrt{\langle f, f \rangle} = \left\{ \int_0^1 f^2(t)dt \right\}^{1/2}.$$

Basis expansions play an important role in FDA. We say that functions $\{e_1, e_2, \dots\}$ form a basis in L^2 if every function $f \in L^2$ admits a unique expansion $f(t) = \sum_{j=1}^{\infty} a_j e_j(t)$. The basis is orthonormal if $\langle e_j, e_{j'} \rangle = 0$ whenever $j \neq j'$ and $\|e_j\| = 1$. In that case, Parseval's identity yields

$$\|f\| = \left\{ \sum_{j=0}^{\infty} a_j^2 \right\}^{1/2}, \quad a_j = \langle f, e_j \rangle.$$

In the statistical FDA framework, the observations are functions. In our setting, we treat the CIDR curve R_n as a realization of a random function R , a random process showing how the return stochastically evolves within a trading day. This means that the norm $\|R\|$, is a random variable. If $E\|R\|^2 < \infty$, we say the random function R is *square integrable*.

The Karhunen-Lo  ve decomposition Consider a random curve $X = \{X(t), t \in [0, 1]\}$ in the space L^2 . We define the mean and covariance functions by $\mu(t) = EX(t)$ and $c(t, s) = E[(X(t) - \mu(t))(X(s) - \mu(s))]$. The Karhunen-Lo  ve decomposition states that every square integrable function X can be represented as

$$X(t) = \mu(t) + \sum_{j=1}^{\infty} \xi_j v_j(t). \quad (3.2)$$

The functions v_j are the eigenfunctions of c , i.e. they are the solutions to the equation

$$\int c(t, s)v(s)ds = \lambda v(t),$$

where the corresponding λ_j are called the eigenvalues of X (or c). They are usually arranged in non-increasing order: $\lambda_1 \geq \lambda_2 \geq \dots$. The random variables ξ_j , called the *scores*, are given by

$$\xi_j = \langle X - \mu, v_j \rangle = \int (X(t) - \mu(t))v_j(t)dt.$$

Decomposition (3.2) is optimal in a sense that X can be well approximated using only very few initial v_j , in many applications 2 or 3 of them. The functions v_j form an orthonormal basis and are

also called the *Functional Principal Components* (FPCs) of X . It can be shown that

$$E\xi_j = 0, \quad E\xi_j^2 = \lambda_j, \quad \text{Cov}(\xi_j, \xi_k) = 0, \quad \text{if } j \neq k$$

and

$$E\|X - \mu\|^2 = \sum_{j=1}^{\infty} \lambda_j. \quad (3.3)$$

The interpretation of the above formula is that λ_j is the variance of the random function X in the principal direction v_j . The sum of these variances is the total variance of X . Relation (3.3) is thus the decomposition of variance corresponding to (3.2).

The above expansions have sample analogs. If X_1, X_2, \dots, X_N is a sample of functions, we define

$$\hat{c}(t, s) = \frac{1}{N} \sum_{i=1}^N (X_i(t) - \hat{\mu}(t))(X_i(s) - \hat{\mu}(s)), \quad \hat{\mu}(t) = \frac{1}{N} \sum_{i=1}^N X_i(t)$$

and

$$\int \hat{c}(t, s) \hat{v}_j(s) ds = \hat{\lambda}_j \hat{v}_j(t). \quad (3.4)$$

If the functions X_i form a stationary, weakly dependent time series in L^2 , the approximation

$$X_i(t) \approx \hat{\mu}(t) + \sum_{j=1}^p \hat{\xi}_{ij} \hat{v}_j(t), \quad \hat{\xi}_{ij} = \langle X_i, \hat{v}_j \rangle,$$

is generally very accurate for small p . In Section 3.3, we will see that for the CIDR curves $p = 2$ or $p = 3$ is sufficient. The above sample approximation allows us to reduce the analysis of statistical properties of curves to the analysis of statistical properties of vectors of reasonably low dimension.

3.2.2 Background on Extreme Value Theory

Peaks-over-Threshold methodology The *peaks-over-threshold* (PoT) analysis of univariate time series has been widely used since the seminal paper of [59]. They advocated the use of the asymptotically motivated Generalized Pareto distribution as a model for the distribution of exceedances over a certain high threshold. Let Y_1, Y_2, \dots, Y_N be a sequence of *scalar* observations with a com-

mon marginal distribution F , $F(y) = P(Y \leq y)$. We are interested in estimating the *conditional excess distribution function*

$$F_u(y) = P(Y \leq u + y \mid Y > u) = \frac{F(u + y) - F(u)}{1 - F(u)}, \quad y > 0,$$

over the threshold u . For large enough u , under mild conditions,

$$F_u(y) \approx G_{\gamma, \sigma_u}(y) := 1 - \left(1 + \frac{\gamma}{\sigma_u} y\right)_+^{-1/\gamma}, \quad (3.5)$$

where γ is called the *shape parameter* and $1/\gamma$ the *tail index*. The distribution G_{γ, σ_u} is called the Generalized Pareto Distribution (GPD). Relation (3.5) can be equivalently stated, for $x > u$, as

$$P(Y > x \mid Y > u) \approx 1 - G_{\gamma, \sigma_u}(x - u) = \left[1 + \gamma \left(\frac{x - u}{\sigma_u}\right)\right]_+^{-1/\gamma}.$$

It follows that

$$P(Y > x) \approx P(Y > u) \left[1 + \gamma \left(\frac{x - u}{\sigma_u}\right)\right]_+^{-1/\gamma}, \quad u \rightarrow \infty.$$

The tail index $1/\gamma$ quantifies the heaviness of the tail. Namely, if $\gamma > 0$, then the asymptotic distribution of the excesses is unbounded and has heavy, Pareto-like tail, $P(Y > x) \sim x^{-1/\gamma}$, $x \rightarrow \infty$. The larger the index γ , the heavier the tail. The case $\gamma < 0$ corresponds to excesses with finite upper bound. Observe that the condition $1 + \gamma(x - u)/\sigma_u \geq 0$ means that $x \leq u + \sigma_u/|\gamma| < \infty$. Finally, the boundary case $\gamma = 0$ corresponds to the exponential distribution since

$$\lim_{\gamma \rightarrow 0} \left[1 + \gamma \left(\frac{x - u}{\sigma_u}\right)\right]_+^{-1/\gamma} = \exp \left\{ -\frac{x - u}{\sigma_u} \right\}, \quad x \geq u.$$

The PoT method involves threshold selection, namely, how large should u be so that the approximation in (3.5) is accurate. Usually, one can assess the stability of parameter estimates, based on fitting of models across a range of threshold values. Alternatively, this can be done by drawing

a mean residual life plot. We produce the mean residual life plot which estimates the function $u \mapsto \mathbb{E}(Y - u \mid Y > u)$. If Y follows a GPD, then the above function should be linear in u . More specifically, if u_0 is a threshold, $\mathbb{E}(Y - u_0 \mid Y > u_0) = \sigma_{u_0}/(1 - \gamma)$ exists, if $\gamma < 1$. Then for any $u > u_0$,

$$\mathbb{E}(Y - u \mid Y > u) = \frac{\sigma_u}{1 - \gamma} = \frac{\sigma_{u_0} + \gamma(u - u_0)}{1 - \gamma}. \quad (3.6)$$

When plotting empirical estimates of the sample means excesses against a range of thresholds, the threshold is chosen to be the lowest level where all the higher threshold based sample mean excesses are consistent with a straight line, once the sample uncertainty is accounted for.

Other traditional ways of choosing threshold involve some values pre-determined by physical considerations or some simple fixed quantile rule, like the upper 10% rule.

Parameter Estimation Having determined a threshold, the parameters can be estimated by the maximum likelihood approach. Suppose that the total number of observations is N , and N_u of them exceed the threshold u . We denote them by the order statistics $x_{(N-N_u+1)}, \dots, x_{(N)}$. Then the log-likelihood derived from (3.5) is

$$l(\sigma_u, \gamma) = -N_u \log \sigma_u - \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^{N_u} \log \left(1 + \frac{x_{(N-i+1)} - u}{\sigma_u} \gamma\right), \quad (3.7)$$

provided $(1 + \sigma^{-1}(x_{(N-i+1)} - u)\gamma) > 0$, for $i = 1, \dots, N_u$. Closed form maximization of the log-likelihood is not possible, so numerical techniques are utilized. Standard errors and confidence intervals for the corresponding parameters are obtained from the likelihood theory. The required asymptotic normality of the maximum likelihood estimators holds for $\gamma > -1/2$. In our context, γ is close to zero.

In practice, the estimation of the parameters γ and σ can be done using the function `gpd.fit` in the R package `ismev`. Estimating $P(Y > u)$ by N_u/N , with the ML estimates $\hat{\gamma}$ and $\hat{\sigma}_u$ obtained from the above procedure leads to

$$\hat{F}(x) \approx 1 - \frac{N_u}{n} \left[1 + \hat{\gamma} \left(\frac{x - u}{\hat{\sigma}_u} \right) \right]_+^{-1/\hat{\gamma}} \quad (3.8)$$

for any $x > u$ such that $1 + \hat{\gamma}(x - u)/\hat{\sigma}_u > 0$.

VaR, return level, and expected shortfall Value-at-Risk (VaR) is one of the most frequently used risk measures in finance. It involves extreme quantile estimation. Suppose a random variable Y with continuous distribution function F models losses or negative returns on an asset over a certain time horizon. Value-at-Risk at the upper tail α , denoted by VaR_α , is defined as the $100(1 - \alpha)$ -th quantile of the distribution F ,

$$\text{VaR}_\alpha(Y) = F^{-1}(1 - \alpha),$$

where F^{-1} is the inverse of the distribution F , and α is typically close to zero. If Y models losses, then $\text{VaR}_\alpha(Y)$ is interpreted as the amount of loss exceeded only $100\alpha\%$ of the time, on the average. A risk measure equivalent to VaR is the *return level*. Let $\alpha = 1/m$, then $L_m = \text{VaR}_{1/m}$ is the *m-period return level*. For instance, it is often more convenient to report return levels on an annual scale. There are about 252 trading days in a year. Taking $m = 252$, we get an upper tail α that approximately equals 0.004 and hence $L_{252} = \text{VaR}_{0.004}$. Thus if Y models a daily loss, then $L_{252}(Y)$ gives us the expected level of loss to be exceeded once a year. The expected shortfall (ES), also called *tail conditional expectation*, is defined as the expected loss that exceeds the VaR_α : $\text{ES}_\alpha = \mathbb{E}(Y \mid Y > \text{VaR}_\alpha)$.

Estimation of the risk measures Denote by $\hat{\sigma}$ and $\hat{\gamma}$ the estimated GPD parameters. To estimate the VaR, we need to estimate $F^{-1}(1 - \alpha)$. This can be done by solving equation (3.8) for x . Then VaR_α can be estimated as

$$\widehat{\text{VaR}}_\alpha = x_\alpha = \begin{cases} u + \frac{\hat{\sigma}}{\hat{\gamma}} \left[\left(\frac{N_u}{N_\alpha} \right)^{\hat{\gamma}} - 1 \right] & \text{if } \hat{\gamma} \neq 0 \\ u + \hat{\sigma} \log \left(\frac{N_u}{N_\alpha} \right) & \text{if } \hat{\gamma} = 0 \end{cases}$$

To estimate the expected shortfall ES_α , we use the relation

$$\widehat{\text{ES}}_\alpha = \widehat{\text{VaR}}_\alpha + \mathbb{E}(Y - \widehat{\text{VaR}}_\alpha \mid Y > \widehat{\text{VaR}}_\alpha),$$

where the second term is the expected exceedance over the threshold $\widehat{\text{VaR}}_\alpha$. Since the *mean excess function* for the GPD with $\gamma < 1$ is given in equation (3.6), it is easy to show that

$$\widehat{\text{ES}}_\alpha = \widehat{\text{VaR}}_\alpha + \frac{\hat{\sigma} + \hat{\gamma} (\widehat{\text{VaR}}_\alpha - u)}{1 - \hat{\gamma}} = \frac{\widehat{\text{VaR}}_\alpha + \hat{\sigma} - \hat{\gamma}u}{1 - \hat{\gamma}}.$$

Large sample standard errors, leading to confidence intervals, for VaR_α and ES_α can be derived using the delta method, see Section 3.4.1.

3.3 Risk analysis of CIDR curves

We now turn to study extremes of the CIDR curves defined in the Introduction. We are specifically interested in the quantification of their extremal behavior before, during and after the financial crisis of 2008. However, the methodology we propose is applicable in many other contexts. A chief question in any analysis of this type is what an *extreme curve* is. Traditional EVT has been developed for samples of scalar observations. Such observations can be ordered, so the concept of “large” observations is clear. Extreme observations are generally those that lie beyond the range of the sample, i.e. have not been observed yet. EVT is then used to compute measures of risk based on upper tail approximations, like those explained in Section 3.2. In Section 3.3.1, we explain what we mean by extreme curves. Then, in Section 3.3.2, we introduce the data sets to which our approaches are applied. Sections 3.3.3 and 3.3.4 include the results of our analysis, which Section 3.3.5 summarizes.

3.3.1 Extreme return curves

We assume that the observed CIDR curves R_n defined in Definition 7 are realizations of a random curve R ; they form a stationary time series in L^2 , with the univariate marginal distribution equal to that of R . The Karhunen-Lo  ve decomposition introduced in Section 3.2 states that the random curve R can be represented as

$$R(t) = \mu(t) + \sum_{j=1}^{\infty} \xi_j v_j(t). \quad (3.9)$$

The mean function μ is assumed to be zero (the estimates $\hat{\mu}(t)$ are not significantly different from zero), so the norm of R is given by

$$\|R\| = \left\{ \int_0^1 R^2(t) dt \right\}^{1/2} = \left\{ \sum_{j=1}^{\infty} \xi_j^2 \right\}^{1/2}, \quad (3.10)$$

where ξ_j are the scores, uncorrelated random variables.

The simplest way to define a “large” CIDR curve is to require that the norm $\|R\|$ be large. By (3.10) this is equivalent to some of the scores being large, irrespective of their sign. There are many ways in which the curve R can be large according to this definition. A very wiggly curve that starts from zero and ends at zero will be large, even though the daily point-to-point return will be zero. Such a situation corresponds to many high frequency scores being large, but the most important first score being close to zero. To explain this point, we show in Figure 3.2 the first three FPCs, v_1, v_2, v_3 for SP500. They look almost identical for all other assets. The first FPC, v_1 , which generally explains about 85% of the variability of these curves, quantifies the upward (if $\xi_1 > 0$) or downward daily trend. This is a pattern that is observed on most trading days. Thus for a zero daily point-to-point return, the coefficient of this component must be close to zero. (The estimated FPC have the same periodicity pattern as those of the Brownian motion, but they rise more steeply at the beginning of the trading day, reflecting the abrupt price changes often seen at the opening of the market.)

Empirical analysis shows that the shapes of the observed CIDR curves are basically encoded by the first three scores:

$$R_n(t) \approx \xi_{n1} v_1(t) + \xi_{n2} v_2(t) + \xi_{n3} v_3(t).$$

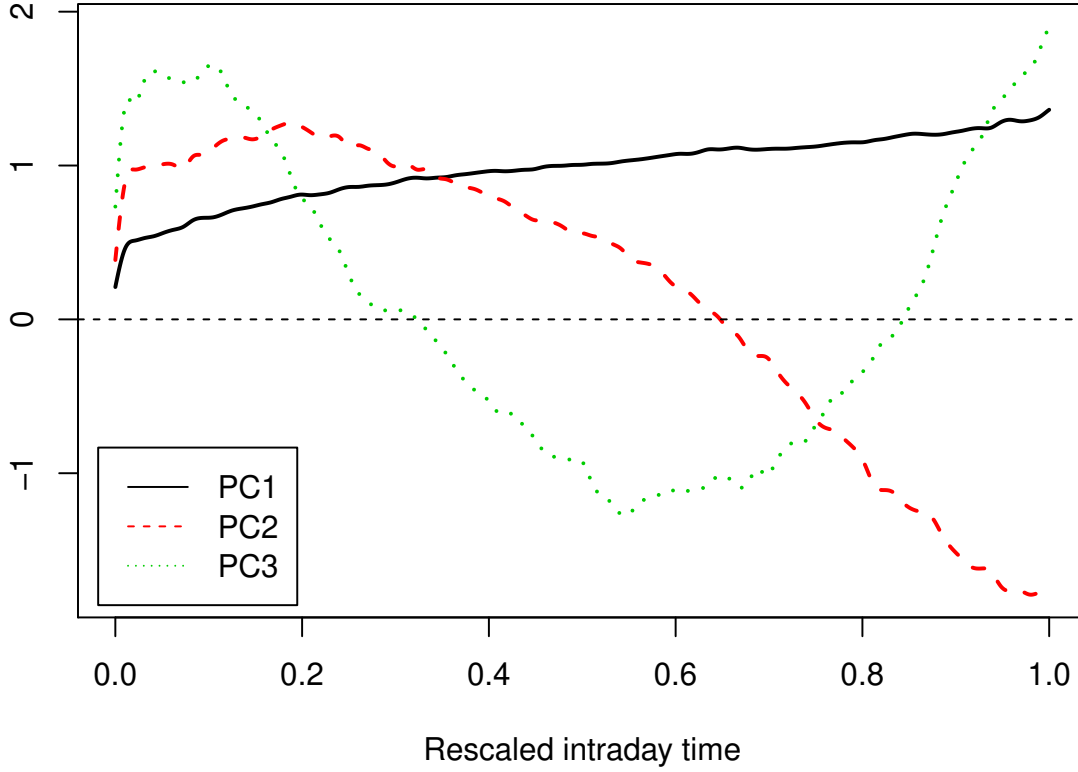


Figure 3.2: The first three FPCs for S&P500 during the period 2006/7/5 - 2011/12/30.

The overwhelming contribution comes from the first and the second FPCs; v_2 quantifies a reversal during a trading day. This is illustrated in Figure 3.3, which shows the plots of $A_n^{(p)} = \{\sum_{j=1}^p \xi_{nj}^2\}^{1/2}$, $p = 2, 3, 4$, against $\|R_n\|$ for a selected asset and 50 trading days. The graphs for other assets and periods are similar. We see that $p = 4$ gives basically the same norm as $p = \infty$, and except for “very small” CIDR curves, using $p = 3$ is sufficient. For days with large market swings, using $p = 2$ is sufficient to describe the intraday market action. Thus, in addition to the norm, it is useful to study the extremal behavior of the sequences $\{\xi_{n1}\}$ and $\{\xi_{n2}\}$. Extremes of these sequences will define extreme curves. In most cases, $\|R\|$ will be large if ξ_1^2 is large, but it does not have to be always the case. The norm $\|R\|$ can be viewed as a new measure of the daily volatility of an asset, but to get a fuller picture of extreme behavior, the first two scores must be considered as well.

To summarize, we will consider a CIDR curve R_n on the trading day n as extreme, if either $\|R_n\|$ is extreme or the scores ξ_{n1} or ξ_{n2} are extreme (or both). The scores can be extreme on either the negative or the positive side.

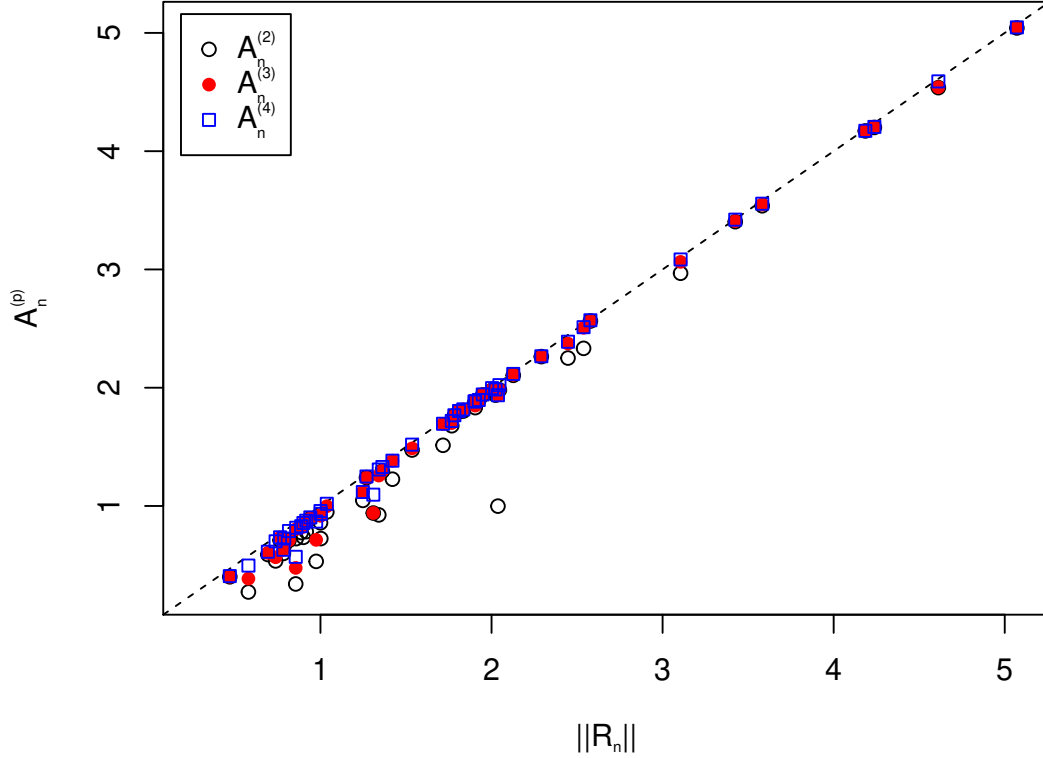


Figure 3.3: The scatter plot of $A_n^{(p)} = \{\sum_{j=1}^p \xi_{nj}^2\}^{1/2}$ against $\|R_n\|$ with $p = 2, 3, 4$ for XLF during the period 2008/6/2 - 2008/8/12 (50 days).

3.3.2 Data description

Our application focuses on nine Select Sector SPDR ETFs, which are proxies for “sector portfolios”. These are Exchange Traded Funds that track the nine S&P 500 sector indexes, they hold individual stocks within the corresponding sectors. For ease of reference, Table 3.1 lists the nine Select Sector SPDR ETFs. The one minute frequency ETF price data are obtained from Quantquote.

Table 3.1: Tickers of the assets (Sector ETFs) used in this study.

Ticker	Sector	Ticker	Sector
XLF	Financials	XLV	Health Care
XLK	Technology	XLI	Industrials
XLY	Consumer Discretionary	XLB	Materials
XLP	Consumer Staples	XLU	Utilities
XLE	Energy		

One of our major objectives is to examine if the extremal behavior of CIDRs for these “portfolios” is “time-varying”. We compare risk measures for periods around the financial crisis. This allows us to determine whether the crisis has a significant impact on the extremal behavior of these assets in a sense defined in Section 3.3.1. We consider four time periods specified in Table 3.2. The inclusion of two time periods after the crisis is justified by our preliminary findings that indicated a difference between the extremal behavior of the cumulative return curves immediately after the generally accepted end of the crisis and at a later time.

Table 3.2: Time periods used in this study.

Designation	Abbreviation	Time span	Sample size (days)
Before	Bf	07/05/2006 – 09/28/2007	313
During	Dr	10/01/2007 – 02/27/2009	351
After 1	A1	03/02/2009 – 07/30/2010	356
After 2	A2	08/02/2010 – 12/30/2011	358

3.3.3 Extremes of the norm and the scores

We first apply the PoT method to the norm (3.10) computed using the sum of the squares of the first 15 scores. The application involves the following steps: select the threshold u , fit a GPD to the exceedances over u , compute point estimates and standard errors of the Value-at-Risk and expected shortfall, see Section 3.2. The results are shown in Table 3.3. The columns “ u ” and “ N_u ”, respectively, show the threshold and the number of observations over that threshold. The column “Parameter $\hat{\gamma}$ ” reports the estimates of the shape parameter with their standard errors. The shape

parameter γ for most assets and periods is seen to be not significantly different from zero, implying that in most cases computation of the risk measures using a Gumbel distribution is appropriate. For longer periods of time, the estimates of γ tend to be positive, which may reflect inhomogeneity of the data or higher risk involved in longer time horizons. To compute the risk measures, we use the GPD with the estimated γ . The last four columns in Table 3.3 show the estimation results for the VaR and ES of $\|R\|$ with upper tail probability $\alpha = 0.004$ and $\alpha = 0.0008$. Since there are 252 trading days, $\text{VaR}_{.004}$ and $\text{VaR}_{.0008}$ are, respectively, the 1-year and 5-year return levels of the norm of the CIDR. The following general pattern for the risk measures for the four periods emerges:

$$\text{during} \geq \text{after1} \geq \text{after2} \approx \text{before}. \quad (3.11)$$

The periods are arranged from the most risky to the least risky. We see that, generally speaking, the impact of the crisis on the extreme risks related to intraday investing persists for about a year and a half after the commonly accepted end of the crisis. Sector specific responses to the crisis are summarized in Figure 3.5. At both levels of VaR and ES, financials (XLF) and Consumer Discretionary (XLY) stocks are seen to be most impacted by the crisis; they reach high levels of risk during the crisis. The XLF stocks continue to be exposed to high risk even one year after the crisis. The Technology (XLK) and Health care (XLV) stocks are seen to be, relatively, least impacted by the crisis. They reach the highest risk during the crisis, and then the risk decays almost linearly with time and eventually (2 or 3 years after the crisis) gets back to the level before the crisis. Among all the ETFs, we also found that the Consumer Staples (XLP) and Utilities (XLU) are the most dramatically affected by the crisis. They both have a fairly low risk levels before and immediately after the crisis, while during the crisis, their risk level drastically goes up to a very high level. The remaining sectors, including Energy (XLE), Industrials (XLI) and Materials (XLB), share a similar risk pattern before, during and after the crisis, and they can be categorized as moderately impacted by the crisis.

Since the first FPC usually contributes about 85% to the total variability in the CIDRs and has a clear interpretation as a monotonic growth or decline during the trading day, we applied our risk

analysis to the magnitude of the first score, $|\xi_1|$. We show the results in Table 3.4. Comparing these results to those in Table 3.3, we see a similar pattern for the risk measures among those 4 periods around the crisis. This suggests that the monotonic trend usually dominates the extreme risk of the CIDRs. The pattern of such gains or losses throughout the trading day can also be summarized by relations (3.11).

We also applied the methodology of Section 3.2.2 to the traditional point-to-point returns, where only the closing prices of a trading day are used to obtain the daily return. Instead of dealing with Left tail (loss) and Right tail (gain) of the returns separately, we only consider the magnitude of the point-to-point returns (or the norms). The corresponding results in Table 3.5 show a different pattern for the risk measures among those 4 periods from that revealed by Table 3.3 and Table 3.4. For the point-to-point returns, the relation can be summarized as follows:

$$\text{during} \geq \text{after2} \geq (\text{or } \approx) \text{after1} \geq \text{before}. \quad (3.12)$$

The only violation of this pattern is XLF, for which “after1” contains more risk than “after2”. Also, another important finding is that for most of the Sector ETFs in the “during” period, the standard errors of VaR and ES are huge, c.f. XLE, XLV and XLU. This indicates that GPD might be an inappropriate model to fit the point-to-point returns in our case. Moreover, this shows us that the proposed *daily volatility* (or norm of CIDRs) reveals different information from that obtained from the traditional point-to-point returns in the sense of extreme value studies. Similarly, sector specific responses to the crisis are summarized in Figure 3.6. We take out two assets XLE and XLU because of their unreasonable standard errors.

Visualization of extremal regions It is of interest to see how the extreme CIDRs might look like. Since in most cases the shape of the CIDRs is sufficiently well explained by the first two FPCs, and in order to keep the graphs informative, we will visualize the extremal regions by focusing on these two scores. We thus consider the approximation

$$R(t) \approx R^{(2)}(t) := \xi_1 v_1(t) + \xi_2 v_2(t).$$

As extreme, we consider $R^{(2)}$ for which $\|R^{(2)}\| = (\xi_1^2 + \xi_2^2)^{1/2}$ is large. We consider extremal regions in four quadrants of the (ξ_1, ξ_2) plane, denoted $(+, +)$, $(+, -)$, $(-, -)$, $(-, +)$, which reflect the four possible combinations of the signs of first two scores. We want to see how these regions change over the “before”, “during”, “after1” and “after2” periods considered above.

Set $\boldsymbol{\xi} = (\xi_1, \xi_2)$. Given a sufficiently large threshold u , one can postulate the following model for the conditional distribution of $\boldsymbol{\xi}$ given $\|\boldsymbol{\xi}\| > u$:

$$\boldsymbol{\xi} \mid \|\boldsymbol{\xi}\| > u \stackrel{d}{\approx} Z(u) \mathbf{Y}(u), \quad (3.13)$$

where $Z(u) \stackrel{d}{\approx} \|\boldsymbol{\xi}\| \mid \{\|\boldsymbol{\xi}\| > u\}$ and $\mathbf{Y}(u) \stackrel{d}{\approx} \boldsymbol{\xi} / \|\boldsymbol{\xi}\| \mid \{\|\boldsymbol{\xi}\| > u\}$ are independent. The random variable $Z(u)$ follows a GPD and $\mathbf{Y}(u)$ follows some probability distribution on the unit sphere. One can fit various models to the data and use them to compute the probabilities $P(\boldsymbol{\xi} \in A)$, where A is outside the ball $B(0, u)$ with a large radius u . This is done using the relation

$$P(\boldsymbol{\xi} \in A) \approx P(\|\boldsymbol{\xi}\| > u) P(Z(u) \mathbf{Y}(u) \in A).$$

The first factor on the right-hand side can be approximated by using standard univariate EVT. The second factor can be calculated by using Monte Carlo methods.

In the Monte Carlo approach, we first simulate realizations of norms from the model for $Z(u)$. Recall that $Z(u)$ follows a GPD $G_u(z)$, which is fitted from the norms $\|\boldsymbol{\xi}_i\| \mid \{\|\boldsymbol{\xi}_i\| > u\}$, where $\{\boldsymbol{\xi}_i = (\xi_{1,i}, \xi_{2,i}), i = 1, \dots, N_u\}$, and N_u is the number of observations with norm greater than u . One can easily draw a new sample of norms $\{z_1, \dots, z_M\}$ from $G_u(z)$, which can be further used to construct new extreme scores. In the second step, we need to simulate data for the angles $\mathbf{Y}(u)$. To this end, we normalize those observed pairs $\boldsymbol{\xi}_i = (\xi_{1,i}, \xi_{2,i})$ whose norm $\|\boldsymbol{\xi}_i\|$ exceeds u to get a pool of angles for the extreme points. That is,

$$\theta_i := \arctan(\xi_{2,i}/\xi_{1,i}),$$

for $i = 1, \dots, N_u$. Based on those observed angles given norms exceeding u , there will be two major ways to generate new angles for future simulation. One could naturally apply a nonparametric method by directly resampling angles from $\{\theta_i, i = 1, \dots, N_u\}$ with replacement. This is just simply a Bootstrap method. Alternatively, we can also employ a kernel density estimator, see Supplementary materials 3.4.2 for the details of this method. From either of these methods, we can simulate a new set of angles $\{\theta_i^*, i = 1, \dots, M\}$. Together with the previously simulated norms $\{z_1, \dots, z_M\}$, we can create simulated score pairs via

$$\boldsymbol{\xi}_i^* = (\xi_1^*, \xi_2^*) = (z_i \cos \theta_i^*, z_i \sin \theta_i^*)$$

with $\|\boldsymbol{\xi}_i^*\| > u$ for $i = 1, \dots, M$. These newly generated scores can be used to depict the extremal region of the CIDRs.

To keep the number of figures relatively small, we merely illustrate our methodology using the XLF (financials) data. The extremal regions obtained using bootstrap are shown in Figures 3.7–3.10. Specifically, in Figure 3.7, we show the extreme region determined by the norms of the CIDRs on the plane of the scores (ξ_1, ξ_2) . The inner black circle on each scatter plot represents the one-year return level ($\text{VaR}_{.004}$) for each of the 4 periods and the bigger grey circle represents a 10,000 day return level ($\text{VaR}_{.0001}$) which can be considered very rarely exceeded. This extremely high level is needed to have an upper limit in the graphs is the time domain. We can see for each of the 4 periods that there are one or two observations, represented by stars, falling into the extreme region defined by $\text{VaR}_{.004}$ and $\text{VaR}_{.0001}$, which approximately represents a region with an upper probability of 0.4%. On the same plot, the little grey circles in the extreme region represent around 100 simulated points. With these simulated points, in Figure 3.8, we show the regions of extreme shapes of the CIDRs for each quadrant and each period corresponding to the extremal regions shown in Figure 3.7. We also plot the real extreme CIDRs, represented by thick solid curves, in the extremal region into which they fall. Similarly, in Figure 3.9 and Figure 3.10, we show the extreme

region relative to the one-year expected shortfall ($ES_{.004}$) on the plane of the scores (ξ_1, ξ_2) , and the corresponding extreme shapes. The expected shortfall is a *coherent* risk measure and constrains both the probability of a tail event and the expected loss given a tail event. This is confirmed in Figure 3.9, where the observed extreme points are very close to the circles that represent norms equal to $ES_{.004}$. Alternatively, as we mentioned above, the procedure of simulating angles for points with norm greater than u can be also done by using the a kernel density estimator (KDE). The corresponding results are shown in Figures 3.11–3.14. A major difference between the two methods in depicting the extreme region can be seen by comparing Figure 3.7 and Figure 3.11, which show the extreme regions of the score pairs using the VaR estimates. With the bootstrap method, we draw samples with angles that behave fairly similarly to the observed data, the KDE method leads to a relatively higher chance of drawing some large angles, which are actually rarely observed in the existing data. For example, a score pair with score 1 approximately equal to score 2 is almost never observed in practice, because the score of the first FPC explains most (about 85%) of the total variability, which leaves very low chance for score 2 to reach a value as high as score 1. Figure 3.8 and Figure 3.12 show that the KDE method leads to a more extended extreme region than the bootstrap method. Similar comments apply to the difference between the two methods for the extreme regions based on the ES.

A general conclusion is that while less intuitive, visualizing extreme curves in the space of scores is more informative, especially if the temporal evolution of these regions is to be examined. Using simple bootstrap may lead to more realistic extreme regions than more sophisticated approaches based on the estimation of the angular density. This requires more investigation which must include suitable criteria.

Fisher’s exact test We examined the independence of signs of (ξ_1, ξ_2) by using Fisher’s exact test, which is a simple contingency table test for independence. We take “XLF” in the “before” period to illustrate, and use $u = 0.8$ as the threshold. The number of norms above u is 52, and the count (proportion) of the points falling in each of the 4 quadrants $(+, +)$, $(+, -)$, $(-, -)$, $(-, +)$ are summarized in the following contingency table:

Frequency (Percentage)	ξ_2		Total
	+	−	
ξ_1 +	15 (0.29)	8 (0.15)	23 (0.44)
−	14 (0.27)	15 (0.29)	29 (0.56)
Total	29 (0.56)	23 (0.44)	52 (1)

Using R function `fisher.test`, one gets the p value of 0.27. The test results for all the assets are summarized in Tables 3.6 and 3.7. For all of them, no evidence was found that there is a preference for any of the four quadrants. Also, no apparent pattern shows up in the 2×2 tables of probabilities as they change over the four periods “before”, “during”, “after1” and “after2”. This leads us to conclude that the crisis has not caused any shift in the “direction” of the extreme CIDR curves. For example, the extreme scores are roughly equally likely to appear in the (+, +) region before and during the crisis.

A more subtle question is if the extreme dependence between the first and second scores has been affected by the crisis. Such an investigation requires rather sophisticated EVT tools, and is the subject of the next section.

3.3.4 Extremal dependence between the magnitude of the scores

By construction, the scores ξ_j in (3.2) are uncorrelated. Using the methods described below, we determined that there is no extremal dependence between the scores at the various levels j . There is however extremal dependence between the squares of level 1 and level 2 scores which occurs for certain assets and certain periods. Roughly speaking, such a dependence means that extremely high monotonic trend increases the chance of a pronounced inflection point, i.e. of a large change in the rate of growth or decline. Independence in this sense means that the occurrence of a pronounced inflection point is not associated with a strong upward or downward trend. This section explores this issue in some detail using measures of extremal dependence defined by [60].

Two measures of extremal dependence Suppose two variables X and Y have a common distribution F . We define

$$\chi = \lim_{z \rightarrow z_+} P(Y > z \mid X > z),$$

where z_+ is the end-point of F , so that $\chi \in [0, 1]$ is a limiting measure of the tendency for one variable to be large conditional on the other one being large. If $\chi = 0$ the variables X and Y are said to be *asymptotically independent*. If $\chi > 0$, they are said to be *asymptotically dependent*.

More generally, suppose that F_X and F_Y are the marginal distribution functions of X and Y , respectively. We then define

$$\begin{aligned}\chi(q) &= P(F_Y(Y) > q \mid F_X(X) > q) \\ &= P(Y > \text{VaR}_{1-q}(Y) \mid X > \text{VaR}_{1-q}(X))\end{aligned}$$

and

$$\chi = \lim_{q \rightarrow 1} \chi(q).$$

The expression for $\chi(q)$ in terms of VaR shows that χ can be interpreted as the probability that Y exceeds its VaR_α value, given that X exceeds its VaR_α value for extremely small α . Thus, if $\chi = 0.5$, for example, 50% of the time the losses of X are due to extreme losses in Y .

Clearly, χ is a useful measure of extremal dependence within the class of asymptotically dependent distributions, where the value of χ increases with the strength of dependence at extreme levels. A complementary measure $\bar{\chi} \in [-1, 1]$ is defined as follows. For $0 < q < 1$, define

$$\bar{\chi}(q) = \frac{2 \log P(F_X(X) > q)}{\log P(F_X(X) > q, F_Y(Y) > q)} - 1$$

and

$$\bar{\chi} = \lim_{q \rightarrow 1} \bar{\chi}(q).$$

When $\bar{\chi} = 1$, the variables X and Y are asymptotically dependent, and when $-1 \leq \bar{\chi} < 1$ they are asymptotically independent. The two measures, χ and $\bar{\chi}$ can thus be used to tests different null hypotheses. Their properties as summarized in the following table, in which independence or dependence refer to these quantities pertaining to extreme events.

Independent	Dependent
$\chi = 0$	$0 < \chi \leq 1$
$-1 \leq \bar{\chi} < 1$	$\bar{\chi} = 1$

Illustrative examples of data models are provided in Section 3.4.3. Generally, extremal independence is more common than extremal dependence. For this reason, one typically first tests the null hypothesis $\bar{\chi} = 1$, and if it is not rejected, assesses the strength of extremal dependence by the value of $\chi > 0$. The details of this procedure are presented in Section 3.4.3.

Application to sector ETFs To illustrate the methodology introduced above, Figure 3.15 shows the plots of $\chi(q)$ and $\bar{\chi}(q)$, together with approximate 95% confidence intervals, for the squared scores ξ_1^2 and ξ_2^2 of XLF. Interpretation of these plots is not completely clear because of the large variance of the estimators. However, we can roughly tell that for the “before” “after1” and “after2” periods $\bar{\chi}(q) \rightarrow 1$ as $q \rightarrow 1$ while for the “during” period this is not the case. In the respective periods, $\chi(q)$ converges to values of around 0.3, 0, 0.3 and 0.4. Furthermore, the likelihood ratio test of asymptotic dependence, see Section 3.4.3, lends support to the use of asymptotically dependent models above a sufficiently high threshold for the “before” “after1” and “after2” periods. Finally, under properly chosen thresholds (we use 80% marginal quantiles as the thresholds for each of the “before” “after1” and “after2” periods), by applying the natural nonparametric estimator (3.19) under the constraint $\hat{\bar{\chi}} = 1$, we obtain $\hat{\chi} = 0.40$ for the “before” period, $\hat{\chi} = 0.43$ for the “after1” period and $\hat{\chi} = 0.35$ for the “after2” period. After accounting for the variance, these results are consistent with the apparent stable levels shown in Figure 3.15.

The estimation results for all the nine ETFs are shown in Table 3.8. For each ETF, the columns “ q ” and “ n_u ”, respectively, give the quantile of the threshold and the number of observations exceeding that threshold during each period. We also give an estimate of $\bar{\chi}$ with corresponding standard errors. The “Pvalue” entries are the results of the likelihood ratio test of asymptotic dependence, with the null hypothesis $\bar{\chi} = 1$. Only if the results are insignificant, do we give a non-zero estimate of χ . The estimation results are sensitive to the choice of the threshold. We choose

it carefully in a multi-stage process. First, we look at the the mean residual life plot using the R function `mrlplot`. Next, we use the R function `tcplot`, which is a plot of parameter estimates at various thresholds for PoT modeling. The valid threshold will be a value such that the parameter estimates depicted are approximately constant above the threshold. We tried several thresholds suggested by the above two approaches and selected a proper quantile q such that the estimation results become stable and reasonable. (An inappropriately chosen q , hence the threshold u , can result in very unstable estimation.)

After accounting for sampling errors, we can clearly observe a pattern among most of the ETFs (XLK, XLY, XLV, XLI and XLB) that the “before” periods show a strong signal of extremal dependence between ξ_1^2 and ξ_2^2 while for some other ETF sectors (XLE and XLU) it occurs in the “after2” period. Since the “after2” is similar to the “before” period in other respects, we can conclude that during the crisis as well as the short period right after the crisis, there tends to be no asymptotic dependence between ξ_1^2 and ξ_2^2 for any ETF sectors. Generally, the extremal dependence is of a different type “before” than in the other three periods, or is of a different type in the “during” period.

3.3.5 Summary and concluding Remarks

We employed the popular extreme value theory *peak-over-threshold* (PoT) method and its extensions to study the extremal behavior of the cumulative intraday returns (CIDRs) of a set of financial assets. We defined and estimated risk measures such as Value-at-Risk (return level) and expected shortfall for function valued data. These and other measures have been applied to the norm and to the scores and their squares. The norm of the CIDRs can be viewed as a new measure of intraday price movements which incorporates intraday variability and systematic trends or changes in direction (e.g. from growth to decline). The scores quantify the presence of specific patterns, like trends or reversals. We proposed several ways of visualizing the risk associated with CIDRs.

Our major findings can be summarized as follows:

1. Scores of the first three FPCs of a CIDR series typically contribute 90% \sim 95% of the total variability. The first score, ξ_1 , typically contributes more than 85%. This indicates that ξ_1 could be a more informative summary of the daily returns than the conventional point-to-point returns based on the closing price. (For example, large ξ_1 indicates that the price increases almost monotonically throughout the day.)
2. The norm of the CIDRs, $\|R\|$, has an exponential-decaying tail when the estimation is performed over a relatively short period (1-2 years). This is likely because the clustering phenomenon is not pronounced over such time periods. The intraday cumulative returns and their variability are either large or small, so the tail, which measures the presence of observations much larger than most observations, is light. This is in contrast to point-to-point returns over long periods of time, which are often modeled as heavy-tailed functions.
3. Estimation of the risk measures leads to the following general pattern around the financial crisis of 2008:

$$\text{during} \geq \text{after1} \geq \text{after2} \approx \text{before},$$

where the periods are arranged from the most risky to the least risky. Since each period has the length of about a year and a half, this implies the crisis had a monotonically decaying, long-term impact on the intraday market risk extending beyond the commonly accepted end of the crisis. The pattern of the decay of risk measures based on point-to-point returns is often not monotonic.

4. Risk measures of CIDRs, in general, contain different information about the extreme behavior of returns than those measures computed for point-to-point returns. The impact of the crisis of 2008 on extreme risk is different for the CIDRs and point-to-point returns, both in terms of a general pattern and impact on specific sectors.
5. The crisis impacted extremal dependence between the volatilities of the principal components of the shapes of the CIDRs.

There is clearly room for improvement and extensions of the methodology that we proposed. Multivariate EVT could in principle be applied to model the joint extremal behavior of the scores, some effort in this direction is outlined in Section 3.4.2, but more sophisticated and accurate approaches might be possible. Rather than using the PoT approach, likelihood based methods that incorporate temporal dependence might be feasible. Potential conditional heteroskedasticity might possibly be incorporated; some research in this direction has been done, [61], [44], [62], but it does not offer any guidance on how to incorporate these models into the study of extremal behavior of CIDR curves. In all, we hope that our exploration will be received with some interest as an attempt to quantify extreme behavior of samples of curves with a potential to quantify some aspects of investment risk.

3.4 Supplementary Materials

3.4.1 Variances of $\widehat{\text{VaR}}_\alpha$ and $\widehat{\text{ES}}_\alpha$

Denote $S_u = P(Y > u)$ and $\hat{S}_u = N_u/N$. Observe that \hat{S}_u is the maximum likelihood estimate of S_u , as the number of exceedances of u follows the Binomial distribution $\sim \text{Bin}(n, S_u)$. Since both $\widehat{\text{VaR}}_\alpha$ and $\widehat{\text{ES}}_\alpha$ can be viewed as a function of $\hat{\sigma}$, $\hat{\gamma}$ and \hat{S}_u , i.e.

$$\widehat{\text{VaR}}_\alpha := g_1(\hat{S}_u, \hat{\sigma}, \hat{\gamma}) = u + \frac{\hat{\sigma}}{\hat{\gamma}} \left[\left(\frac{\hat{S}_u}{\alpha} \right)^{\hat{\gamma}} - 1 \right]$$

and

$$\widehat{\text{ES}}_\alpha := g_2(\hat{S}_u, \hat{\sigma}, \hat{\gamma}) = u + \frac{\hat{\sigma}}{1 - \hat{\gamma}} + \frac{\hat{\sigma}}{\hat{\gamma}(1 - \hat{\gamma})} \left[\left(\frac{\hat{S}_u}{\alpha} \right)^{\hat{\gamma}} - 1 \right],$$

large sample standard errors or confidence intervals for VaR_α and ES_α can be derived from the delta method. From standard properties of the binomial distribution, $\text{VaR}(\hat{S}_u) \approx \hat{S}_u(1 - \hat{S}_u)/n$ and denote the (i, j) term of the variance-covariance matrix of $\hat{\sigma}$ and $\hat{\gamma}$ by $v_{i,j}$, then the complete covariance-variance matrix for $(\hat{\sigma}, \hat{\gamma}, \hat{S}_u)$ is approximately

$$V = \begin{bmatrix} \hat{S}_u(1 - \hat{S}_u)/n & 0 & 0 \\ 0 & v_{1,1} & v_{1,2} \\ 0 & v_{2,1} & v_{2,2} \end{bmatrix}.$$

Hence, by the delta method,

$$\text{Var}(\widehat{\text{VaR}}_\alpha) \approx \nabla g_1^T V \nabla g_1$$

and

$$\text{Var}(\widehat{\text{ES}}_\alpha) \approx \nabla g_2^T V \nabla g_2,$$

where

$$\begin{aligned} \nabla g_1^T &= \left[\frac{\partial g_1}{\partial S_u}, \frac{\partial g_1}{\partial \sigma}, \frac{\partial g_1}{\partial \gamma} \right] \\ &= \left\{ \frac{\sigma S_u^{\gamma-1}}{\alpha^\gamma}, \frac{1}{\gamma} \left[\left(\frac{S_u}{\alpha} \right)^\gamma - 1 \right], -\frac{\sigma}{\gamma^2} \left[\left(\frac{S_u}{\alpha} \right)^\gamma - 1 \right] + \frac{\sigma}{\gamma} \left(\frac{S_u}{\alpha} \right)^\gamma \log \left(\frac{S_u}{\alpha} \right) \right\} \end{aligned}$$

and

$$\begin{aligned} \nabla g_2^T &= \left[\frac{\partial g_2}{\partial S_u}, \frac{\partial g_2}{\partial \sigma}, \frac{\partial g_2}{\partial \gamma} \right] \\ &= \left\{ \frac{\sigma S_u^{\gamma-1}}{(1-\gamma)\alpha^\gamma}, \frac{1}{\gamma(1-\gamma)} \left[\gamma + \left(\frac{S_u}{\alpha} \right)^\gamma - 1 \right], \right. \\ &\quad \left. \frac{\sigma}{(1-\gamma)^2} + \frac{\sigma(2\gamma-1)}{\gamma^2(1-\gamma)^2} \left[\left(\frac{S_u}{\alpha} \right)^\gamma - 1 \right] + \frac{\sigma}{\gamma(1-\gamma)} \left(\frac{S_u}{\alpha} \right)^\gamma \log \left(\frac{S_u}{\sigma} \right) \right\} \end{aligned}$$

evaluated at $(\hat{S}_u, \hat{\sigma}, \hat{\gamma})$.

3.4.2 Estimation of the angular density

This section explains how the density of $\mathbf{Y}(u)$ in (3.13) can be estimated nonparametrically. Since we consider two scores, this is a density on $(-\pi, \pi)$. The idea is that if $\theta_i, i = 1, \dots, n$, are *i.i.d.* points from a density f , then its kernel estimator is given by

$$\hat{f}_h(\theta) := \frac{2}{n} \sum_{i=1}^n \phi_h(\tan(\theta - \theta_i)/2) / \cos^2((\theta - \theta_i)/2), \quad (3.14)$$

where $\phi_h(x) = \phi(x/h)/h$ is probability density on the real line. The idea behind this estimator is to lift a distribution on the real line to the unit circle via the transformation

$$\Theta \rightarrow \tan(\Theta/2).$$

By default, we will use the standard normal density for ϕ (see also [63]). To apply this idea to the procedure of resampling angles, suppose $(\xi_{1,i}, \xi_{2,i}), i = 1, \dots, N_u$ are the points with norm greater than the threshold u and define

$$\theta_i := \arctan(\xi_{2,i}/\xi_{1,i})$$

for $i = 1, \dots, N_u$. Choosing the bandwidth $h = 2\pi/N_u^{2/3}$, we compute (3.14) using the θ_i 's, and so obtain an estimated density function $\hat{f}_h(\theta)$ of the angles θ on $(-\pi, \pi)$. From this density, we sample angles that can be used to generate new extreme points. Figure 3.4 illustrates this method using the asset XLF in the “before” period.

3.4.3 Elaboration on the measures χ and $\bar{\chi}$

This section provides details of inference based on the extreme dependence measures χ and $\bar{\chi}$. Essentially, χ provides a measure with which to describe the strength of dependence within the class of asymptotically dependent variables, while $\bar{\chi}$ provides a corresponding measure within the class of asymptotically independent variables. Taken together, the pair $(\chi, \bar{\chi})$ provides more complete information characterizing the form and the degree of extremal dependence of any two random variables. For asymptotically dependent variables, $\bar{\chi} = 1$ and the value of $\chi > 0$ measures the strength of dependence. For asymptotically independent variables, $\chi = 0$, and the value of $\bar{\chi} \in [-1, 1)$ might be used to quantify the strength of extremal independence, which may be interpreted also as a secondary or hidden tail dependence in terms of the notion of *hidden regular variation* [64].

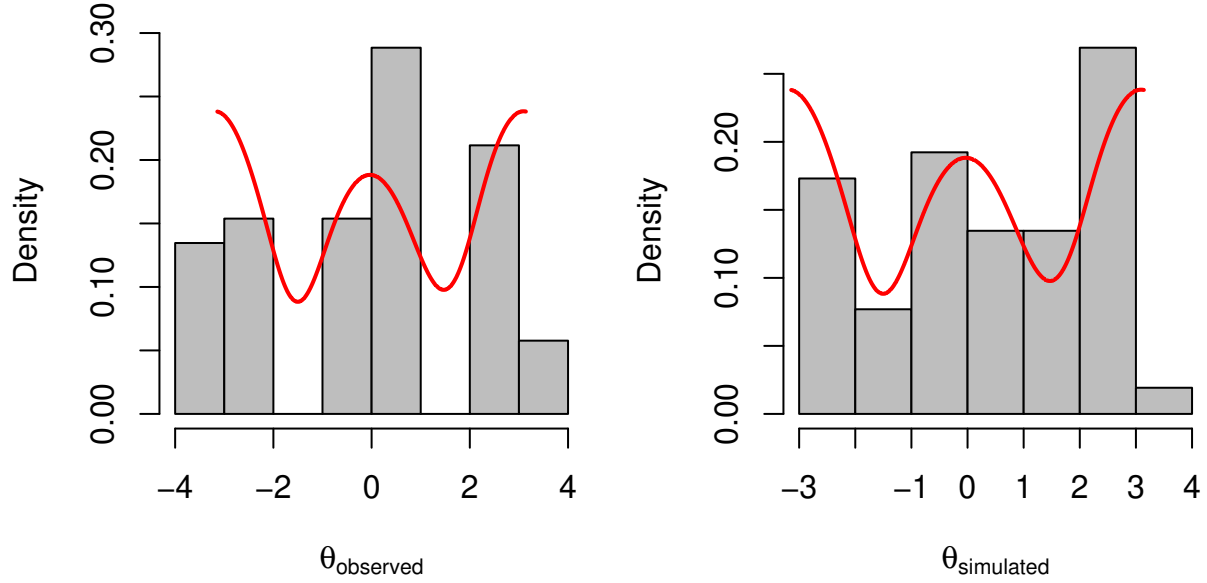


Figure 3.4: Left panel: histogram of the observed θ_i 's with the estimate $\hat{f}_h(\theta)$ superposed. Right panel: histogram of the same number of new θ_i 's generated from $\hat{f}_h(\theta)$.

We begin by providing two examples where $\chi > 0$ (asymptotic dependence) and $\chi = 0$ (asymptotic independence).

Example 3.4.1. In this example, X and Y have a common risk factor. Let Z , ε_x and ε_y be independent $\text{GPD}(\gamma, \gamma, 0)$ random variables with the same index $\gamma > 0$, such that

$$\mathbb{P}(Z > u) = \mathbb{P}(\varepsilon_x > u) = \mathbb{P}(\varepsilon_y > u) = (1 + u)^{-1/\gamma},$$

for $u > 0$. Consider the simple factor model

$$X = \sigma_z Z + a\varepsilon_x \quad \text{and} \quad Y = \sigma_z Z + b\varepsilon_y,$$

where σ_z , a and b are positive. By applying Lemma 2 below with $a_1 = b_1 = \sigma_z$, $a_2 = a$, $b_2 = 0$, and $a_3 = 0$, $b_3 = b$, we obtain

$$\sigma_X^{1/\gamma} = \sigma_z^{1/\gamma} + a^{1/\gamma}, \quad \sigma_Y^{1/\gamma} = \sigma_z^{1/\gamma} + b^{1/\gamma},$$

$$\chi(X, Y) = \frac{\sigma_z^{1/\gamma}}{\sigma_z^{1/\gamma} + a^{1/\gamma}} \wedge \frac{\sigma_z^{1/\gamma}}{\sigma_z^{1/\gamma} + b^{1/\gamma}} = \frac{1}{1 + \max(a^{1/\gamma}, b^{1/\gamma})/\sigma_z^{1/\gamma}}.$$

Thus, the larger the ratio $\max(a^{1/\gamma}, b^{1/\gamma})/\sigma_z^{1/\gamma}$, the smaller the extremal dependence. This is natural since in this case, σ_z is relatively small and the common factor $\sigma_z Z$ has relatively smaller contribution to the simultaneous extremes of X and Y . On the other hand, if σ_z dominates a and b , the common risk factor leads to more frequent simultaneous extremes, which is reflected by relatively larger values of the extremal dependence measure χ .

Example 3.4.2. Suppose now that (X, Y) are jointly Gaussian with correlation coefficient $\rho \in (-1, 1)$. It can be shown that $\chi = 0$ in this case, i.e. X and Y are always asymptotically independent (unless $\rho = 1$). The measure $\bar{\chi}$ is more appropriate in this case and $\bar{\chi} = \rho$. See [65] for the details.

Lemma 2. Let Z_1, \dots, Z_p be independent heavy-tailed $\text{GPD}(\gamma, \gamma, 0)$ random variables, i.e., $\mathbb{P}(Z_i > u) = (1 + u)^{-1/\gamma}$, $u > 0$, for some $\gamma > 0$. Define

$$X = \sum_{i=1}^p a_i Z_i \quad \text{and} \quad Y = \sum_{i=1}^p b_i Z_i,$$

where $a_i > 0$ and $b_j > 0$, for some $i, j \in \{1, \dots, p\}$. Then,

$$\mathbb{P}(X > u, Y > u) \sim u^{-1/\gamma} \sum_{i=1}^p (a_i \wedge b_i)_+^{1/\gamma}, \quad \text{as } u \rightarrow \infty, \quad (3.15)$$

where $x_n \sim y_n$ means $x_n/y_n \rightarrow 1$, $x \wedge y := \min\{x, y\}$ and $(x)_+ := \max\{x, 0\}$.

Consequently, the tail dependence coefficient between X and Y equals

$$\chi(X, Y) = \sum_{i=1}^p \left(\frac{a_i}{\sigma_X} \wedge \frac{b_i}{\sigma_Y} \right)_+^{1/\gamma}, \quad (3.16)$$

where $\sigma_X^{1/\gamma} = \sum_{i=1}^p (a_i)_+^{1/\gamma}$ and $\sigma_Y^{1/\gamma} = \sum_{i=1}^p (b_i)_+^{1/\gamma}$.

Proof. We shall first obtain Relation (3.16) as a consequence of (3.15). For simplicity, let $\alpha := 1/\gamma > 0$. By formally applying (3.15) to $a_i := b_i$, we obtain

$$\mathbb{P}(Y > u) \sim \sigma_Y^\alpha u^{-\alpha}, \text{ where } \sigma_X^\alpha = \sum_{i=1}^p (b_i)_+^\alpha,$$

and similarly $\mathbb{P}(X > u) \sim \sigma_X^\alpha u^{-\alpha}$, as $u \rightarrow \infty$. Observe that $\sigma_X > 0$ and $\sigma_Y > 0$ since $\max_{i=1,\dots,p} a_i > 0$ and $\max_{i=1,\dots,p} b_i > 0$.

Thus, $\mathbb{P}(X/\sigma_X > u) \sim \mathbb{P}(Y/\sigma_Y > u) \sim u^{-\alpha}$, as $u \rightarrow \infty$, and hence

$$\chi(X, Y) = \lim_{u \rightarrow \infty} \mathbb{P}(X/\sigma_X > u | Y/\sigma_Y > u) = \lim_{u \rightarrow \infty} u^\alpha \mathbb{P}(X > \sigma_X u, Y > \sigma_Y u).$$

Now, by applying (3.15) with a_i replaced by a_i/σ_X and b_i replaced by b_i/σ_Y , the formula (3.16) follows.

Relation (3.15) is well known but we give a proof for completeness. This result can be understood in terms of the so-called *one big jump* heuristic. Namely, since the Z_i 's are independent and heavy-tailed with the same tail exponent $\alpha = 1/\gamma > 0$, asymptotically, a linear combination of the Z_i 's is large if one and only one of its terms is extreme. That is, the probabilities that two or more components $a_i Z_i$ contribute to an extreme value of X is asymptotically negligible. Using this principle, one can intuitively see that

$$\mathbb{P}(X > u, Y > u) \sim \sum_{i=1}^p \mathbb{P}(a_i Z_i > u, b_i Z_i > u) = u^{-\alpha} \sum_{i=1}^p (a_i \wedge b_i)_+^\alpha,$$

as $u \rightarrow \infty$.

We will make the above heuristic precise using the notion of multivariate regular variation. We start with some terminology. A set $A \subset \mathbb{R}^p \setminus \{\mathbf{0}\}$ is said to be bounded away from $\mathbf{0}$ if there is a ball $B(\mathbf{0}, \epsilon)$ centered at the origin with radius $\epsilon > 0$ such that $A \cap B(\mathbf{0}, \epsilon) = \emptyset$. A random vector $\mathbf{Z} = (Z_1, \dots, Z_p)^\top$ is said to be *regularly varying* in $\mathbb{R}^p \setminus \{\mathbf{0}\}$ if there exists a Borel measure ν on $\mathbb{R}^p \setminus \{\mathbf{0}\}$, such that

$$c(u)\mathbb{P}(\mathbf{Z} \in uA) \rightarrow \nu(A), \quad (3.17)$$

for all measurable bounded away from $\mathbf{0}$ and ν -continuity sets A , i.e., such that $\nu(\partial A) = 0$, where $\partial A = \overline{A} \setminus A^\circ$ is the boundary of A .

In the simple case above, it is easy to see that the vector \mathbf{Z} with independent $\text{GPD}(\gamma, \gamma, 0)$ components is regularly varying, where ν is supported on the positive orthant $[0, \infty)^p$. One can identify the measure ν by taking $A = [\mathbf{0}, \mathbf{x}]^c = \{\mathbf{y} \in \mathbb{R}^p : y_i > x_i \text{ for some } i = 1, \dots, p\}$, for $\mathbf{x} \geq \mathbf{0}$.

Indeed, by using the independence of the Z_i 's and the inclusion-exclusion formula, we get as $u \rightarrow \infty$,

$$\mathbb{P}(\mathbf{Z} \in uA) \sim \mathbb{P}(\cup_{i=1}^p \{Z_i > ux_i\}) \sim u^{-\alpha} \sum_{i=1}^p x_i^\alpha =: u^{-\alpha} \nu(A).$$

Thus (3.17) holds with $c(u) := u^\alpha$, where the measure ν is concentrated on the positive axes $\ell_i := \{\lambda \mathbf{e}_i : \lambda > 0\}$, $i = 1, \dots, p$, with $\mathbf{e}_1 = (1, 0, \dots, 0)^\top, \dots, \mathbf{e}_p = (0, \dots, 0, 1)^\top$ denoting the standard basis of \mathbb{R}^p . In general, for a measurable set A , we have the formula

$$\nu(A) = \sum_{i=1}^p \nu_\alpha(\pi_i(A \cap \ell_i)), \quad (3.18)$$

where $\nu_\alpha(x, \infty) = x^{-\alpha}$, $x > 0$ is a measure on the positive half-line $(0, \infty)$, and $\pi_i : \mathbb{R}^p \rightarrow \mathbb{R}$ is the projection on the i -th coordinate axis.

With this general tool, we can establish the tail behavior of various functionals of \mathbf{Z} by relating them to suitable sets A . In particular, observe that

$$\{X > u\} = \{\mathbf{Z} \in uA\}, \quad \text{and} \quad \{Y > u\} = \{\mathbf{Z} \in uB\}.$$

where $A := \{\mathbf{z} = (z_i)_{i=1}^p : \sum_{i=1}^p a_i z_i > 1\}$ and $B := \{\mathbf{z} : \sum_{i=1}^p b_i z_i > 1\}$. Observe that both A and B are bounded away from $\mathbf{0}$ and it can be shown, using the scaling properties of the measure ν , that A and B are both ν -continuity sets. Thus, by (3.17) applied to the set $A \cap B$, we obtain

$$u^\alpha \mathbb{P}(X > u, Y > u) = u^\alpha \mathbb{P}(\mathbf{Z} \in u(A \cap B)) \rightarrow \nu(A \cap B),$$

as $u \rightarrow \infty$. By elementary geometric considerations, however, $(A \cap B) \cap \ell_i = \emptyset$, unless both a_i and b_i are positive. In this case, if $a_i \wedge b_i > 0$, we have

$$\pi_i(A \cap B \cap \ell_i) = (1/a_i, \infty) \cap (1/b_i, \infty) = (1/(a_i \wedge b_i), \infty).$$

This, in view of (3.18), yields the formula (3.15) and completes the proof. \square

Equivalent definition of χ Before we discuss any estimation methods of χ and $\bar{\chi}$, it is useful to introduce an alternative way of defining the two measures. The joint distribution of a set of variables can be separated into their respective marginal distributions and dependence structure among them. This idea is also well known in the study of *copulas*. In order to focus on the dependence structure of two variables, it is helpful to remove the influence of marginal aspects first by transforming the raw data to a common marginal distribution. After such a transformation, differences in distributions are purely due to dependence structures. Now we transform the bivariate variables (X, Y) to unit Fréchet marginals S and T as follows:

$$S = -1/\log F_X(X) \quad \text{and} \quad T = -1/\log F_Y(Y),$$

where F_X and F_Y are the marginal distribution functions of X and Y respectively. Notice that there are two typical ways to estimate F_X and F_Y practically. Each of the two marginals can either be approximated from a GPD family using univariate extreme value theory, or be simply approximated by its empirical cumulative distribution function (ECDF). It follows that S and T have the common distribution function $F(s) = e^{-1/s}$. Thus, we can show that χ can be equivalently defined as

$$\begin{aligned}
\chi &= \lim_{q \rightarrow 1} P(F(T) > q \mid F(S) > q) \\
&= \lim_{s \rightarrow \infty} P(T > s \mid S > s) \\
&= \lim_{s \rightarrow \infty} \frac{P(T > s, S > s)}{P(S > s)}.
\end{aligned}$$

Estimation of χ and $\bar{\chi}$ Ledford and Tawn ([66], [67]) established that under weak conditions

$$P(S > s, T > s) \sim \mathcal{L}(s)s^{-1/\eta} \text{ as } s \rightarrow \infty,$$

where $0 < \eta \leq 1$ is a constant and \mathcal{L} is a slowly varying function. From this representation we have

$$\bar{\chi} = 2\eta - 1,$$

and if $\bar{\chi} = 1$, corresponding to $\eta = 1$, then $\chi = \lim_{s \rightarrow \infty} \mathcal{L}(s)$. So the estimation of η and $\lim_{s \rightarrow \infty} \mathcal{L}(s)$ provide the basis for estimating χ and $\bar{\chi}$. Here, the key point is to estimate the joint distribution $P(S > s, T > s)$ hence η and $\mathcal{L}(s)$. Let $Z = \min(S, T)$, so $P(S > s, T > s) = P(\min(S, T) > s) = P(Z > s)$. Inference follows using univariate extreme value techniques to fit a generalized Pareto distribution to the data points in Z that exceed a large fixed threshold u , then the estimated shape parameter of the fitted distribution provides an estimate of η .

Before we estimate χ , it is important to decide if there exists an asymptotic dependence. We thus first test the null hypothesis $\bar{\chi} = 1$. Only if there is no significant evidence to reject it, we estimate χ .

Testing $\bar{\chi} = 1$ There are basically two major different ways to test if $\bar{\chi} = 1$. First, we can simply create a confidence interval for an estimate of $\bar{\chi}$. From the discussion above, the estimate and standard error of the shape parameter η can be obtained in the usual way from standard likelihood theory. So by the Delta method, one can obtain the corresponding estimate and standard error of $\bar{\chi}$ as

$$\hat{\bar{\chi}} = 2\hat{\eta} - 1 \quad \text{and} \quad SE(\hat{\bar{\chi}}) = 2SE(\hat{\eta}),$$

where $SE(\hat{\eta})$ can be obtained from the standard maximum likelihood method. Hence a 95% confidence interval of the true $\bar{\chi}$ is approximated by

$$\left(\hat{\bar{\chi}} - 1.96SE(\hat{\bar{\chi}}), \hat{\bar{\chi}} + 1.96SE(\hat{\bar{\chi}}) \right),$$

as the sample size in the study is big enough. Thus, the hypothesis $\bar{\chi} = 1$ is rejected (at level 0.05) when the confidence interval does not capture 1. For the second way, given that $\bar{\chi} = 1$ corresponds to $\eta = 1$, one can use `anova` in R to perform a likelihood ratio test for asymptotic dependence, with the null hypothesis $\eta = 1$ versus the alternative $\eta < 1$. We refer to [68] for the details of R implementation.

Estimation of χ Once the above test shows no significant evidence to reject $\bar{\chi} = 1$, we can estimate χ under the assumption that $\bar{\chi} = \eta = 1$. [69] provided a maximum-likelihood estimator of χ , i.e. the natural non-parametric estimator under the constraint $\hat{\bar{\chi}} = 1$

$$\hat{\chi} = \frac{un_u}{n} \tag{3.19}$$

and

$$SE(\hat{\chi}) = \sqrt{\frac{u^2 n_u (n - n_u)}{n^3}},$$

where n is the sample size and n_u is the number of observations of variable Z that exceed the threshold u .

3.4.4 Large tables and displays

In this section, we display all the estimation results of the risk analysis of CIDR curves using large Tables 3.3–3.8 as well as Figures 3.5–3.15.

Table 3.3: Estimation results for the norm $\|R\|$ for nine ETFs.

In this table, the Value-at-Risk estimates $\text{VaR}_{.004}$ and $\text{VaR}_{.0008}$, respectively, represent the 1-year and 5-years return levels. Standard errors in parentheses. Recall that u is the threshold and N_u is the sample size of values over the threshold.

ETF	Prd	u	N_u	Parameter $\hat{\gamma}$	$\text{VaR}_{.004}$	$\text{ES}_{.004}$	$\text{VaR}_{.0008}$	$\text{ES}_{.0008}$
XLF (Financials)	Bf	0.8	54	0.12(0.16)	2.46(0.16)	3.08(0.13)	3.43(0.94)	4.18(0.34)
	Dr	2.7	71	-0.27(0.09)	7.83(0.20)	8.40(0.70)	8.78(0.43)	9.15(0.97)
	A1	2.5	33	-0.24(0.20)	6.81(0.37)	7.54(1.57)	8.02(1.23)	8.52(2.56)
	A2	1.3	51	0.02(0.16)	3.11(0.13)	3.65(0.19)	3.98(0.69)	4.54(0.35)
XLK (Technology)	Bf	0.8	67	-0.08(0.11)	2.13(0.04)	2.39(0.09)	2.56(0.13)	2.79(0.17)
	Dr	1.6	69	-0.08(0.11)	4.96(0.23)	5.64(0.43)	6.06(0.79)	6.66(0.74)
	A1	1.1	73	0.16(0.12)	3.08(0.19)	3.90(0.19)	4.33(1.09)	5.39(0.65)
	A2	0.8	68	-0.10(0.15)	1.99(0.03)	2.23(0.11)	2.37(0.14)	2.57(0.21)
XLY (Cons. Dis.)	Bf	0.6	78	0.09(0.11)	1.91(0.06)	2.33(0.05)	2.58(0.31)	3.06(0.12)
	Dr	2.2	58	-0.02(0.13)	7.06(0.69)	8.29(1.11)	9.04(2.92)	10.22(2.03)
	A1	1.4	63	0.10(0.17)	4.20(0.41)	5.18(0.33)	5.74(2.49)	6.88(0.64)
	A2	0.9	69	-0.06(0.15)	2.56(0.08)	2.92(0.16)	3.14(0.35)	3.47(0.31)
XLP (Cons. Stap.)	Bf	0.6	50	-0.15(0.15)	1.34(0.01)	1.47(0.08)	1.55(0.04)	1.65(0.15)
	Dr	0.6	185	0.20(0.08)	5.38(0.82)	7.27(0.43)	8.25(4.41)	10.84(1.37)
	A1	0.8	72	-0.10(0.12)	1.92(0.02)	2.13(0.07)	2.27(0.09)	2.45(0.14)
	A2	0.6	66	0.12(0.14)	1.51(0.04)	1.84(0.06)	2.03(0.21)	2.43(0.21)
XLE (Energy)	Bf	1.3	50	-0.08(0.19)	2.46(0.05)	2.71(0.16)	2.87(0.23)	3.09(0.32)
	Dr	2.5	51	-0.23(0.14)	5.79(0.16)	6.27(0.60)	6.58(0.49)	6.91(0.94)
	A1	1.7	51	-0.23(0.15)	3.45(0.05)	3.71(0.23)	3.88(0.14)	4.06(0.39)
	A2	1.3	57	-0.12(0.13)	2.93(0.05)	3.24(0.16)	3.44(0.19)	3.70(0.30)
XLV (Health)	Bf	0.5	66	0.11(0.14)	1.74(0.09)	2.24(0.18)	2.51(0.54)	3.16(0.77)
	Dr	1.5	53	-0.06(0.17)	4.46(0.29)	5.15(0.58)	5.59(1.31)	6.22(1.08)
	A1	0.8	95	0.12(0.11)	3.13(0.20)	3.93(0.13)	4.38(1.01)	5.35(0.30)
	A2	0.7	71	0.11(0.11)	2.20(0.08)	2.73(0.08)	3.03(0.41)	3.66(0.19)
XLI (Industrials)	Bf	0.8	48	-0.12(0.16)	1.69(0.02)	1.87(0.10)	1.98(0.08)	2.12(0.20)
	Dr	1.5	88	0.14(0.13)	5.28(0.68)	6.67(0.41)	7.44(3.92)	9.17(0.88)
	A1	1.5	65	-0.10(0.11)	3.83(0.10)	4.28(0.22)	4.57(0.31)	4.95(0.39)
	A2	1.2	48	-0.10(0.17)	2.42(0.04)	2.69(0.14)	2.85(0.17)	3.08(0.28)
XLB (Materials)	Bf	0.8	85	0.09(0.13)	2.75(0.16)	3.35(0.09)	3.70(0.81)	4.39(0.18)
	Dr	2.0	74	0.02(0.14)	5.83(0.51)	6.87(0.57)	7.49(2.40)	8.57(1.07)
	A1	1.5	70	-0.05(0.14)	3.99(0.17)	4.54(0.30)	4.88(0.71)	5.38(0.54)
	A2	1.2	64	-0.23(0.10)	2.84(0.03)	3.06(0.13)	3.21(0.07)	3.36(0.21)
XLU (Utilities)	Bf	0.8	57	-0.01(0.15)	2.17(0.07)	2.52(0.10)	2.74(0.32)	3.08(0.20)
	Dr	0.9	163	0.13(0.09)	5.78(0.75)	7.35(0.34)	8.22(3.53)	10.13(0.67)
	A1	1.1	56	-0.06(0.14)	2.83(0.08)	3.24(0.17)	3.49(0.32)	3.86(0.32)
	A2	0.7	74	0.16(0.14)	2.17(0.12)	2.77(0.17)	3.10(0.74)	3.88(0.73)

Table 3.4: Estimation results for the magnitude of the first score $|\xi_1|$ for nine ETFs.

Quantities reported in this table are the same as in Table 3.3.

ETF	Prd	u	N_u	Parameter $\hat{\gamma}$	VaR. _{.004}	ES. _{.004}	VaR. _{.0008}	ES. _{.0008}
XLF (Financials)	Bf	0.7	60	0.10(0.14)	2.30(0.12)	2.86(0.10)	3.17(0.64)	3.82(0.21)
	Dr	2.5	69	-0.30(0.10)	7.57(0.18)	8.11(0.69)	8.46(0.38)	8.79(0.95)
	A1	2.3	34	-0.23(0.19)	6.72(0.40)	7.49(1.58)	8.00(1.35)	8.54(2.60)
	A2	1.2	57	0.02(0.15)	2.98(0.12)	3.48(0.16)	3.78(0.58)	4.30(0.30)
XLK (Technology)	Bf	0.9	49	-0.07(0.15)	2.07(0.04)	2.34(0.10)	2.50(0.16)	2.74(0.20)
	Dr	1.6	60	-0.06(0.13)	4.91(0.29)	5.66(0.52)	6.13(1.14)	6.82(0.94)
	A1	1.0	76	0.14(0.12)	3.05(0.18)	3.84(0.16)	4.27(0.99)	5.25(0.43)
	A2	0.9	45	-0.03(0.23)	1.97(0.06)	2.26(0.13)	2.44(0.33)	2.72(0.27)
XLY (Cons. Dis.)	Bf	0.6	68	0.12(0.13)	1.87(0.08)	2.32(0.07)	2.58(0.40)	3.12(0.22)
	Dr	2.2	51	-0.04(0.14)	6.97(0.64)	8.14(1.17)	8.86(2.56)	9.95(2.13)
	A1	1.0	108	0.06(0.11)	4.04(0.28)	4.88(0.21)	5.37(1.23)	6.30(0.37)
	A2	1.0	48	-0.14(0.17)	2.45(0.05)	2.74(0.20)	2.92(0.21)	3.14(0.39)
XLP (Cons. Stap.)	Bf	0.6	44	-0.19(0.14)	1.31(0.01)	1.43(0.07)	1.50(0.03)	1.59(0.15)
	Dr	0.5	188	0.19(0.08)	5.22(0.77)	7.06(0.40)	8.03(4.06)	10.54(1.24)
	A1	0.9	48	-0.07(0.15)	1.89(0.03)	2.13(0.08)	2.28(0.12)	2.49(0.17)
	A2	0.5	90	0.13(0.12)	1.47(0.04)	1.81(0.07)	2.01(0.21)	2.43(0.29)
XLE (Energy)	Bf	1.2	52	-0.21(0.15)	2.39(0.02)	2.57(0.15)	2.68(0.08)	2.80(0.26)
	Dr	2.5	45	-0.14(0.17)	5.62(0.24)	6.24(0.71)	6.64(0.96)	7.14(1.26)
	A1	1.7	46	-0.23(0.16)	3.40(0.05)	3.66(0.24)	3.83(0.14)	4.01(0.42)
	A2	1.2	57	-0.17(0.12)	2.85(0.04)	3.13(0.15)	3.31(0.13)	3.52(0.27)
XLV (Health)	Bf	0.5	58	0.15(0.16)	1.69(0.09)	2.17(0.16)	2.43(0.53)	3.05(0.66)
	Dr	1.7	35	-0.18(0.22)	4.31(0.18)	4.81(0.80)	5.14(0.79)	5.51(1.45)
	A1	0.7	104	0.10(0.10)	3.07(0.18)	3.83(0.12)	4.26(0.84)	5.16(0.24)
	A2	0.7	59	0.08(0.11)	2.16(0.08)	2.66(0.08)	2.94(0.35)	3.51(0.16)
XLI (Industrials)	Bf	0.8	39	-0.25(0.15)	1.64(0.01)	1.76(0.10)	1.84(0.03)	1.93(0.19)
	Dr	1.3	88	0.11(0.12)	5.14(0.60)	6.43(0.42)	7.16(3.15)	8.70(0.82)
	A1	1.6	53	-0.07(0.13)	3.80(0.12)	4.30(0.26)	4.62(0.45)	5.07(0.48)
	A2	1.2	41	-0.09(0.21)	2.37(0.05)	2.65(0.18)	2.82(0.24)	3.06(0.38)
XLB (Materials)	Bf	0.8	76	0.13(0.15)	2.74(0.21)	3.44(0.13)	3.83(1.20)	4.68(0.38)
	Dr	1.5	94	-0.04(0.11)	5.65(0.39)	6.52(0.52)	7.06(1.51)	7.86(0.91)
	A1	1.6	48	-0.22(0.11)	3.94(0.07)	4.31(0.27)	4.55(0.17)	4.81(0.43)
	A2	1.3	50	-0.20(0.12)	2.79(0.03)	3.03(0.15)	3.19(0.09)	3.36(0.26)
XLU (Utilities)	Bf	0.8	50	0.06(0.19)	2.16(0.11)	2.60(0.09)	2.86(0.62)	3.34(0.16)
	Dr	0.9	136	0.06(0.08)	5.35(0.48)	6.56(0.34)	7.26(1.97)	8.60(0.61)
	A1	1.1	50	-0.01(0.16)	2.80(0.10)	3.27(0.18)	3.55(0.48)	4.00(0.35)
	A2	0.6	93	0.11(0.12)	2.04(0.08)	2.53(0.06)	2.81(0.41)	3.39(0.17)

Table 3.5: Estimation results for the magnitude of the point-to-point returns for nine ETFs.

Quantities reported in this table are the same as in Table 3.3.

ETF	Prd	u	N_u	Parameter $\hat{\gamma}$	VaR. _{.004}	ES. _{.004}	VaR. _{.0008}	ES. _{.0008}
XLF (Financials)	Bf	1.1	59	0.06(0.18)	4.58(0.65)	5.72(0.47)	6.28(3.61)	7.53(0.97)
	Dr	5.3	53	-0.23(0.13)	15.53(1.37)	16.00(2.77)	17.95(3.82)	17.96(4.27)
	A1	4.9	26	-0.21(0.23)	13.52(2.03)	14.46(4.77)	16.47(7.46)	16.89(8.80)
	A2	2.4	53	-0.02(0.14)	7.77(0.90)	9.15(1.08)	10.06(3.94)	11.40(2.16)
XLK (Technology)	Bf	1.4	40	-0.03(0.14)	2.83(0.07)	3.16(0.08)	3.44(0.25)	3.76(0.17)
	Dr	2.9	58	0.10(0.14)	9.20(1.77)	11.75(1.59)	12.72(9.59)	15.66(3.46)
	A1	2.1	46	-0.10(0.12)	5.05(0.19)	5.51(0.30)	6.11(0.61)	6.48(0.57)
	A2	1.8	48	0.02(0.18)	4.89(0.41)	5.84(0.40)	6.37(2.17)	7.34(0.85)
XLY (Cons. Dis.)	Bf	1.1	57	-0.04(0.15)	3.26(0.15)	3.71(0.17)	4.06(0.61)	4.48(0.35)
	Dr	3.4	50	-0.04(0.14)	9.50(1.08)	10.90(1.33)	11.98(4.38)	13.29(2.61)
	A1	2.5	43	-0.15(0.12)	5.95(0.23)	6.31(0.42)	7.07(0.69)	7.28(0.76)
	A2	2.0	42	-0.05(0.16)	5.35(0.35)	6.11(0.50)	6.75(1.49)	7.43(1.02)
XLP (Cons. Stap.)	Bf	0.8	51	0.04(0.14)	2.27(0.09)	2.76(0.08)	2.99(0.41)	3.50(0.16)
	Dr	1.8	56	0.11(0.15)	6.65(1.11)	8.64(1.01)	9.43(6.23)	11.75(2.19)
	A1	1.2	60	-0.10(0.11)	3.24(0.08)	3.54(0.13)	3.91(0.24)	4.16(0.24)
	A2	1.2	42	0.00(0.16)	3.25(0.16)	3.87(0.19)	4.24(0.78)	4.86(0.41)
XLE (Energy)	Bf	1.7	67	-0.17(0.17)	3.82(0.11)	3.89(0.17)	4.34(0.41)	4.33(0.35)
	Dr	3.5	64	0.04(0.14)	15.68(5.60)	19.38(5.57)	21.36(27.57)	25.29(11.57)
	A1	2.7	51	-0.12(0.10)	6.16(0.21)	6.57(0.34)	7.30(0.58)	7.59(0.60)
	A2	2.3	57	-0.03(0.12)	6.90(0.57)	8.00(0.68)	8.77(2.24)	9.82(1.32)
XLV (Health)	Bf	0.7	64	-0.12(0.15)	2.53(0.08)	2.78(0.17)	3.06(0.31)	3.25(0.32)
	Dr	2.2	47	0.06(0.16)	7.84(1.47)	9.91(1.56)	10.88(7.99)	13.16(3.41)
	A1	1.6	38	-0.16(0.14)	3.98(0.12)	4.25(0.26)	4.77(0.37)	4.93(0.50)
	A2	1.4	57	-0.02(0.14)	4.44(0.29)	5.17(0.33)	5.68(1.26)	6.38(0.67)
XLI (Industrials)	Bf	1.2	46	-0.10(0.18)	3.04(0.11)	3.33(0.17)	3.68(0.45)	3.92(0.36)
	Dr	2.9	55	-0.17(0.12)	8.21(0.45)	8.67(0.81)	9.67(1.35)	9.92(1.36)
	A1	2.3	58	-0.10(0.13)	6.42(0.35)	7.03(0.51)	7.77(1.25)	8.25(0.94)
	A2	2.2	45	-0.12(0.14)	5.87(0.31)	6.41(0.52)	7.13(1.09)	7.55(0.98)
XLB (Materials)	Bf	1.3	59	-0.19(0.13)	3.72(0.09)	3.87(0.21)	4.31(0.28)	4.36(0.37)
	Dr	4.1	40	-0.22(0.14)	10.59(0.70)	10.93(1.48)	12.37(1.98)	12.39(2.46)
	A1	2.6	58	-0.32(0.12)	5.51(0.07)	5.19(0.20)	6.03(0.16)	5.58(0.32)
	A2	2.6	40	-0.09(0.19)	6.66(0.53)	7.40(0.77)	8.22(2.40)	8.83(1.60)
XLU (Utilities)	Bf	1.3	43	-0.13(0.21)	3.80(0.21)	4.15(0.39)	4.62(0.95)	4.88(0.82)
	Dr	2.4	51	0.11(0.17)	9.60(2.84)	12.66(2.66)	13.91(17.54)	17.51(6.12)
	A1	1.7	51	-0.25(0.15)	3.59(0.05)	3.52(0.14)	4.03(0.16)	3.86(0.27)
	A2	1.1	70	0.19(0.14)	3.90(0.44)	5.40(0.58)	5.81(2.87)	7.75(1.78)

Table 3.6: Fisher's exact test (with frequencies).

This table shows the results of Fisher's exact test on independence of the four regions with frequencies of points falling in each region given their norms above the threshold.

ETF	Prd	u	N_u	(+, +)	(+, -)	(-, -)	(-, +)	P value
XLF (Financials)	Bf	0.8	52	15	8	15	14	0.27
	Dr	2.7	68	13	21	18	16	0.62
	A1	2.5	29	6	7	6	10	0.47
	A2	1.3	50	14	10	12	14	0.78
XLK (Technology)	Bf	0.8	61	14	12	15	20	1.00
	Dr	1.6	64	18	15	15	16	1.00
	A1	1.1	69	16	14	22	17	0.47
	A2	0.8	66	16	13	20	17	0.62
XLY (Cons. Dis.)	Bf	0.6	74	19	16	18	21	1.00
	Dr	2.2	53	12	16	12	13	0.59
	A1	1.4	61	18	12	16	15	0.44
	A2	0.9	63	18	15	10	20	0.44
XLP (Cons. Stap.)	Bf	0.6	47	13	8	13	13	0.56
	Dr	0.6	171	45	48	37	41	0.65
	A1	0.8	67	14	17	16	20	0.47
	A2	0.6	60	15	12	14	19	1.00
XLE (Energy)	Bf	1.3	46	13	12	8	13	0.56
	Dr	2.5	48	12	11	8	17	0.38
	A1	1.7	50	12	15	13	10	1.00
	A2	1.3	54	16	8	15	15	0.27
XLV (Health)	Bf	0.5	63	15	17	16	15	1.00
	Dr	1.5	51	16	8	13	14	0.39
	A1	0.8	92	25	19	29	19	0.14
	A2	0.7	65	17	10	19	19	0.32
XLI (Industrials)	Bf	0.8	44	10	11	12	11	1.00
	Dr	1.5	77	21	18	17	21	1.00
	A1	1.5	63	16	15	17	15	0.80
	A2	1.2	46	10	11	14	11	1.00
XLB (Materials)	Bf	0.8	81	21	19	20	21	1.00
	Dr	2.0	67	19	12	11	25	0.61
	A1	1.5	60	13	15	17	15	1.00
	A2	1.2	61	20	11	13	17	0.61
XLU (Utilities)	Bf	0.8	55	15	12	15	13	0.59
	Dr	0.9	147	39	38	34	36	1.00
	A1	1.1	53	11	13	14	15	0.78
	A2	0.7	69	20	18	13	18	0.81

Table 3.7: Fisher's exact test (with probabilities).

This table shows the results of Fisher's exact test on independence of the four regions with probabilities of points falling in each region given their norms above the threshold.

ETF	Prd	u	N_u	P(+, +)	P(+, -)	P(-, -)	P(-, +)	P value
XLF (Financials)	Bf	0.8	52	0.29	0.15	0.29	0.27	0.27
	Dr	2.7	68	0.19	0.31	0.26	0.24	0.62
	A1	2.5	29	0.21	0.24	0.21	0.34	0.47
	A2	1.3	50	0.28	0.20	0.24	0.28	0.78
XLK (Technology)	Bf	0.8	61	0.23	0.20	0.25	0.33	1.00
	Dr	1.6	64	0.28	0.23	0.23	0.25	1.00
	A1	1.1	69	0.23	0.20	0.32	0.25	0.47
	A2	0.8	66	0.24	0.20	0.30	0.26	0.62
XLY (Cons. Dis.)	Bf	0.6	74	0.26	0.22	0.24	0.28	1.00
	Dr	2.2	53	0.23	0.30	0.23	0.25	0.59
	A1	1.4	61	0.30	0.20	0.26	0.25	0.44
	A2	0.9	63	0.29	0.24	0.16	0.32	0.44
XLP (Cons. Stap.)	Bf	0.6	47	0.28	0.17	0.28	0.28	0.56
	Dr	0.6	171	0.26	0.28	0.22	0.24	0.65
	A1	0.8	67	0.21	0.25	0.24	0.30	0.47
	A2	0.6	60	0.25	0.20	0.23	0.32	1.00
XLE (Energy)	Bf	1.3	46	0.28	0.26	0.17	0.28	0.56
	Dr	2.5	48	0.25	0.23	0.17	0.35	0.38
	A1	1.7	50	0.24	0.30	0.26	0.20	1.00
	A2	1.3	54	0.30	0.15	0.28	0.28	0.27
XLV (Health)	Bf	0.5	63	0.24	0.27	0.25	0.24	1.00
	Dr	1.5	51	0.31	0.16	0.25	0.27	0.39
	A1	0.8	92	0.27	0.21	0.32	0.21	0.14
	A2	0.7	65	0.26	0.15	0.29	0.29	0.32
XLI (Industrials)	Bf	0.8	44	0.23	0.25	0.27	0.25	1.00
	Dr	1.5	77	0.27	0.23	0.22	0.27	1.00
	A1	1.5	63	0.25	0.24	0.27	0.24	0.80
	A2	1.2	46	0.22	0.24	0.30	0.24	1.00
XLB (Materials)	Bf	0.8	81	0.26	0.23	0.25	0.26	1.00
	Dr	2.0	67	0.28	0.18	0.16	0.37	0.61
	A1	1.5	60	0.22	0.25	0.28	0.25	1.00
	A2	1.2	61	0.33	0.18	0.21	0.28	0.61
XLU (Utilities)	Bf	0.8	55	0.27	0.22	0.27	0.24	0.59
	Dr	0.9	147	0.27	0.26	0.23	0.24	1.00
	A1	1.1	53	0.21	0.25	0.26	0.28	0.78
	A2	0.7	69	0.29	0.26	0.19	0.26	0.81

Table 3.8: Estimation results of extreme dependence between ξ_1^2 and ξ_2^2 for the nine ETFs.

Standard errors in parentheses. The value of q is the quantile for the threshold. Pvalue is for the likelihood ratio test of $\bar{\chi} = 1$, and “*” indicates cases where $\bar{\chi}$ is not significantly different from 1.

ETF	Prd	q	n_u	$\hat{\bar{\chi}}$	Pvalue	$\hat{\chi}$
XLF (Financials)	Bf	0.80	63	0.59(0.43)	0.21*	0.40(0.04)
	Dr	0.70	106	0.25(0.32)	0.00	0
	A1	0.80	71	0.53(0.42)	0.15*	0.43(0.05)
	A2	0.80	72	0.73(0.41)	0.39*	0.35(0.04)
XLK (Technology)	Bf	0.80	63	0.67(0.43)	0.32*	0.38(0.04)
	Dr	0.70	105	0.01(0.33)	0.00	0
	A1	0.80	71	-0.20(0.35)	0.00	0
	A2	0.75	89	0.21(0.33)	0.00	0
XLY (Cons. Dis.)	Bf	0.70	94	0.62(0.34)	0.15*	0.42(0.04)
	Dr	0.75	87	0.28(0.39)	0.02	0
	A1	0.65	125	0.31(0.28)	0.00	0
	A2	0.75	90	0.37(0.37)	0.04	0
XLP (Cons. Stap.)	Bf	0.80	63	0.29(0.41)	0.04	0
	Dr	0.60	140	0.62(0.33)	0.13*	0.42(0.03)
	A1	0.85	54	-0.03(0.43)	0.01	0
	A2	0.75	90	0.42(0.38)	0.05	0
XLE (Energy)	Bf	0.80	63	-0.59(0.32)	0.00	0
	Dr	0.80	70	0.15(0.40)	0.01	0
	A1	0.80	71	0.21(0.37)	0.01	0
	A2	0.70	108	0.61(0.35)	0.14*	0.39(0.03)
XLV (Health)	Bf	0.80	63	0.97(0.50)	0.93*	0.38(0.04)
	Dr	0.80	70	-0.10(0.33)	0.00	0
	A1	0.70	106	0.25(0.30)	0.00	0
	A2	0.85	54	0.19(0.42)	0.02	0
XLI (Industrials)	Bf	0.85	47	0.54(0.47)	0.22*	0.37(0.05)
	Dr	0.80	70	-0.58(0.28)	0.00	0
	A1	0.80	70	-0.16(0.35)	0.00	0
	A2	0.80	72	0.23(0.39)	0.02	0
XLB (Materials)	Bf	0.85	47	0.44(0.50)	0.16*	0.35(0.05)
	Dr	0.80	70	-0.37(0.34)	0.00	0
	A1	0.80	71	-0.36(0.30)	0.00	0
	A2	0.80	72	0.13(0.36)	0.01	0
XLU (Utilities)	Bf	0.80	62	0.07(0.37)	0.00	0
	Dr	0.80	70	0.09(0.38)	0.01	0
	A1	0.80	71	0.31(0.40)	0.04	0
	A2	0.80	72	0.87(0.45)	0.70*	0.36(0.04)

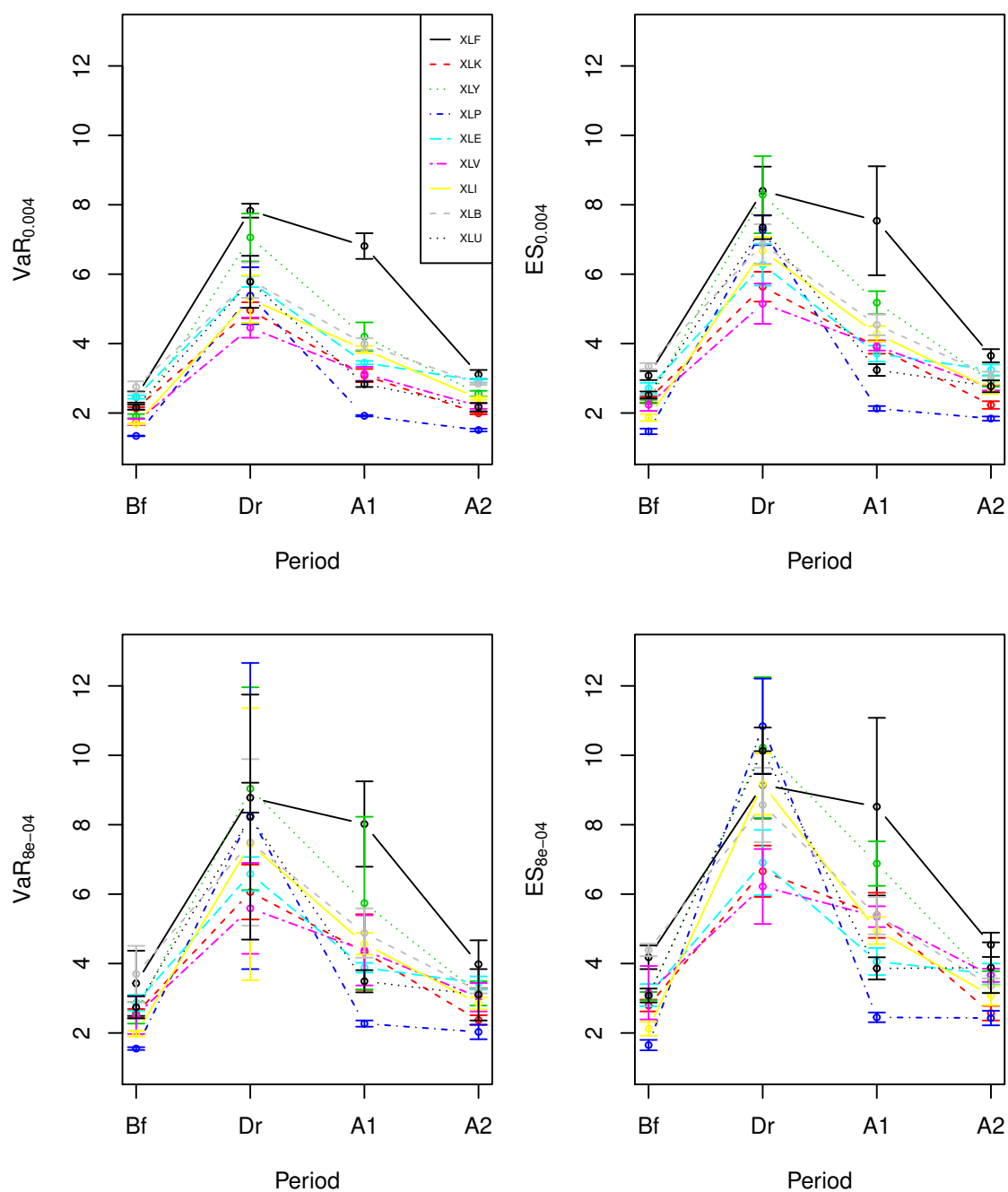


Figure 3.5: VaR & ES Estimates of norms of **CIDRs**. This is an illustration of the VaR & ES Estimates (with standard errors) of norms of CIDRs for the nine Sector EFTs.

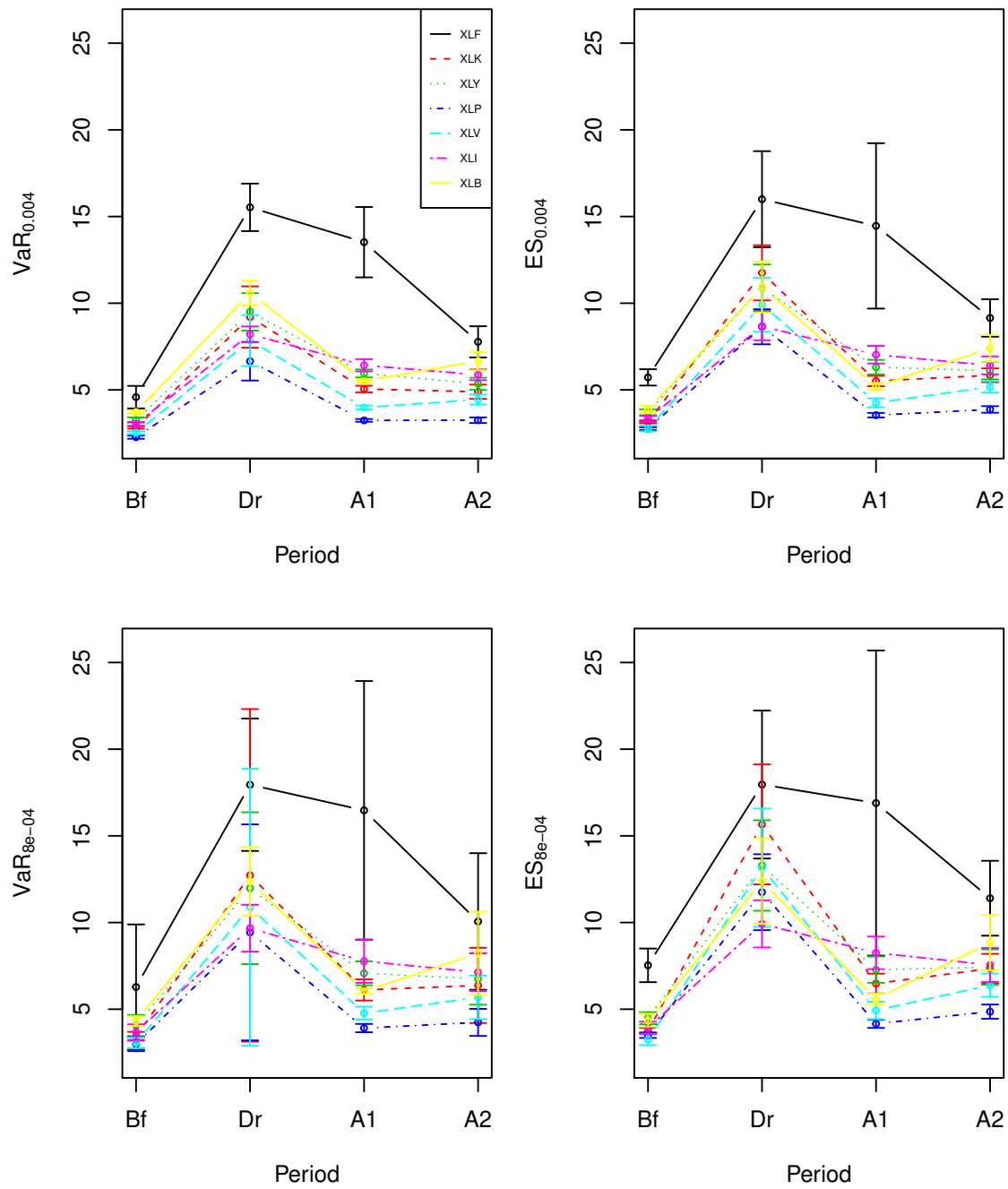


Figure 3.6: VaR & ES Estimates of magnitude of **point-to-point** returns. This is an illustration of the VaR & ES Estimates (with standard errors) of magnitude of point-to-point returns for the seven Sector EFTs.

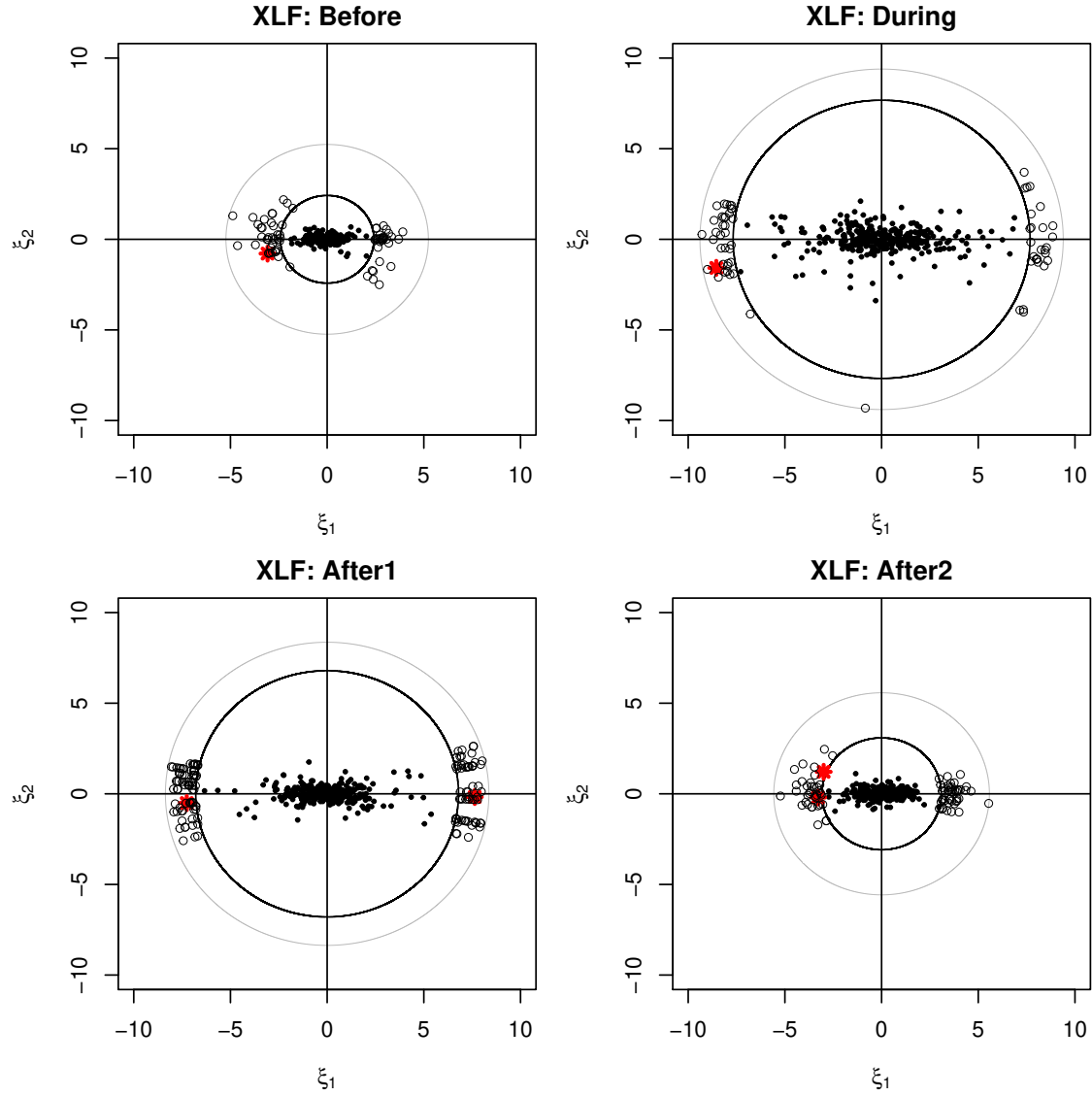


Figure 3.7: The scatter plot of XLF scores with **VaR Estimates** (Bootstrap). On each of the scatter plot, the big black thick circle corresponds to the norm at $\text{VaR}_{0.004}$ (or the 1-year return level), while the bigger grey circle corresponds to the norm at $\text{VaR}_{0.0001}$. The stars represent the real extreme points with norm greater than $\text{VaR}_{0.004}$. The little grey circles represent the simulated extreme points with norm greater than $\text{VaR}_{0.004}$ based on the Bootstrap method.

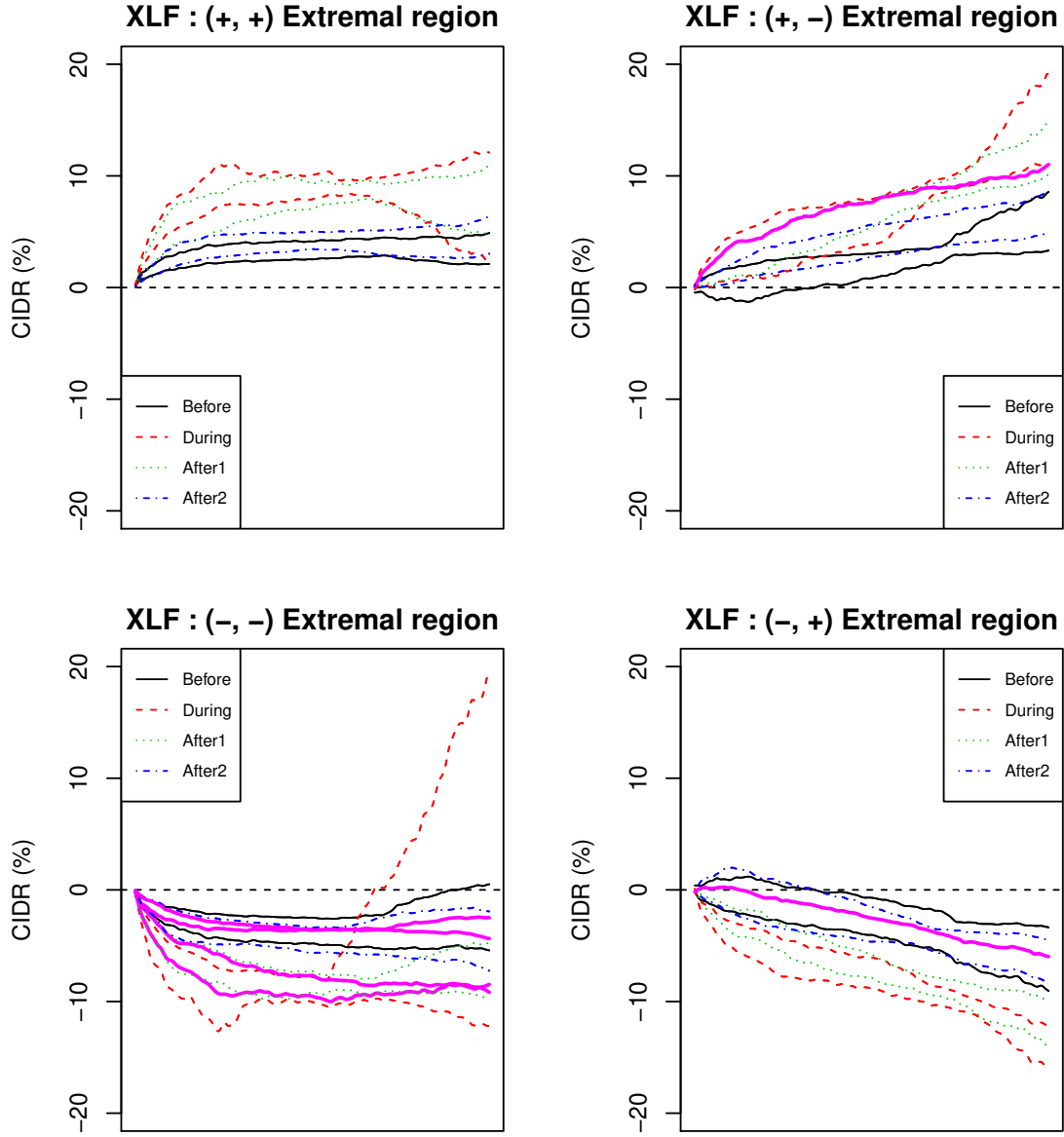


Figure 3.8: VaR Extremal regions of XLF (Bootstrap). This is the extremal regions of CIDRs corresponding to extreme regions in different quadrants on the (ξ_1, ξ_2) plane in Figure 3.7. On each panel, the two curves in the same type, respectively, represents the lower and upper bound of the extremal region of CIDRs regarding one of the four periods under comparison. The thick solid curves represent the real extremal CIDRs with norm greater than $\text{VaR}_{0.004}$.

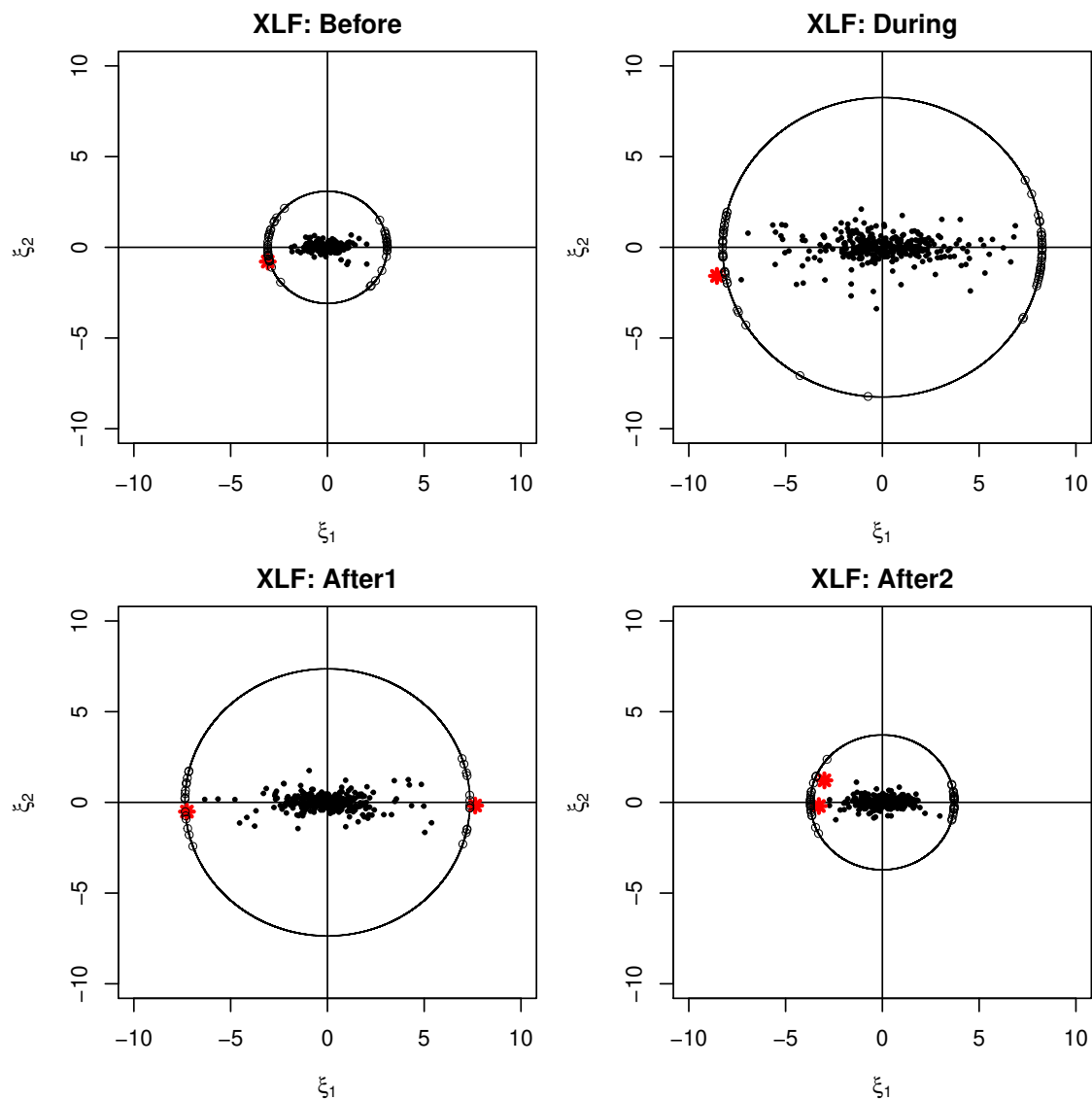


Figure 3.9: The scatter plot of XLF scores with **ES Estimates** (Bootstrap). On each of the scatter plot, the big black thick circle corresponds to the norm at $ES_{0.004}$ (or the 1-year expected shortfall). The stars represent the real extreme points with norms greater than $VaR_{0.004}$. The little grey circles represent the simulated extreme points with norm at $ES_{0.004}$ based on the Bootstrap method.

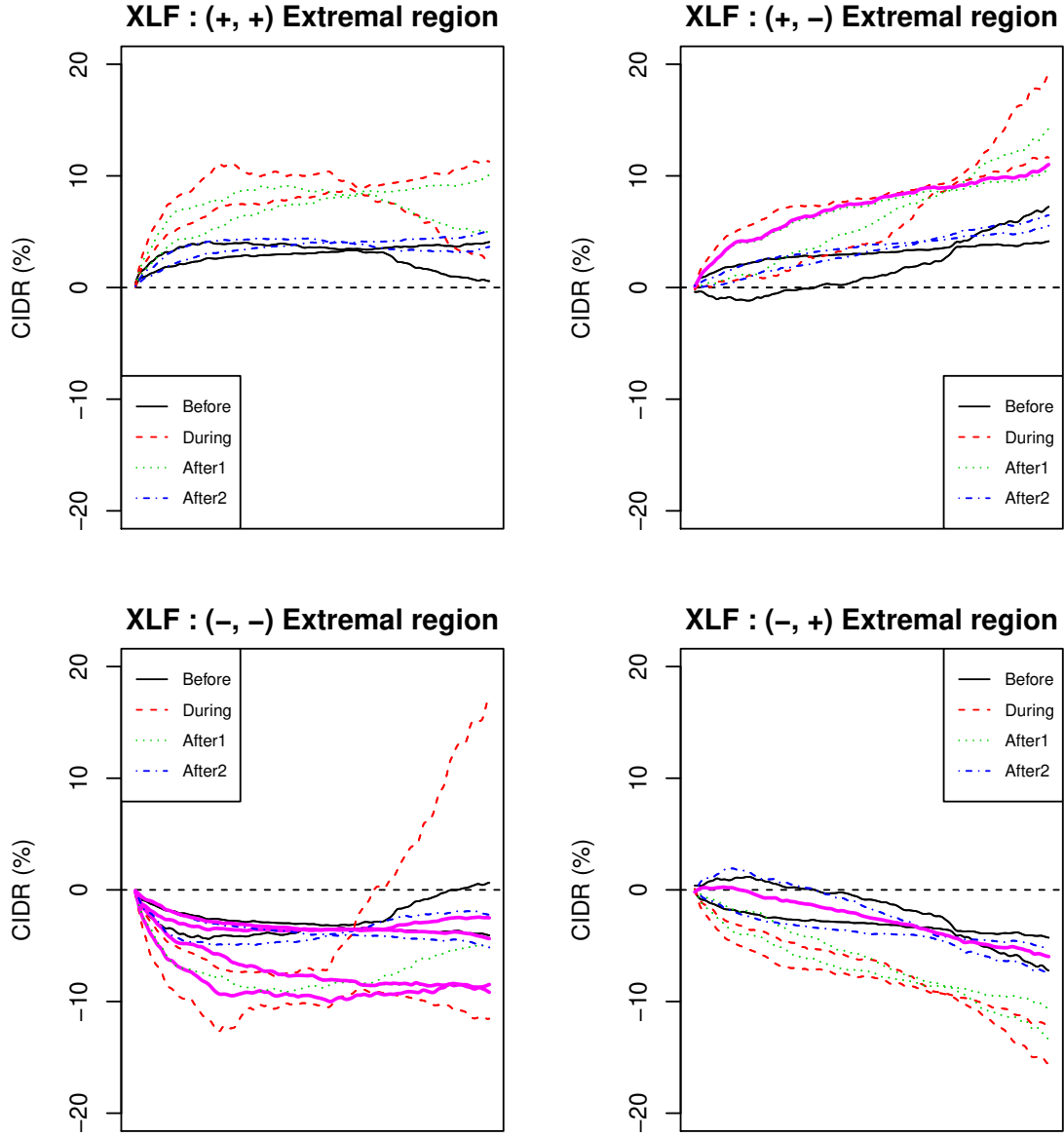


Figure 3.10: ES Extremal regions of XLF (Bootstrap). This is the extremal regions of CIDRs corresponding to extreme regions in different quadrants on the (ξ_1, ξ_2) plane in Figure 3.9. On each panel, the two curves in the same type, respectively, represents the lower and upper bound of the extremal region of CIDRs regarding one of the four periods under comparison. The thick solid curves represent the real extremal CIDRs with norm greater than $VaR_{0.004}$.

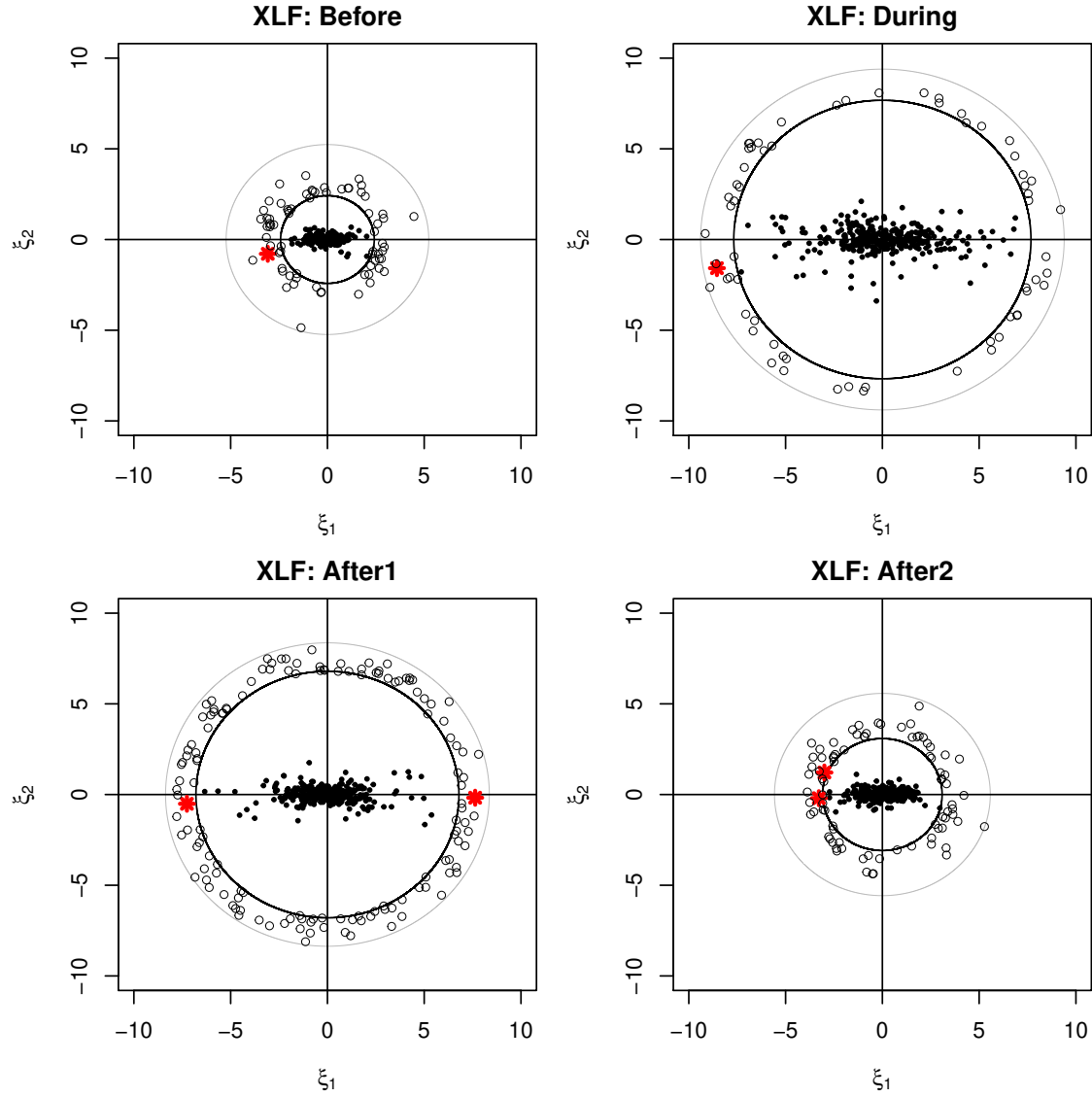


Figure 3.11: The scatter plot of XLF scores with **VaR Estimates** (KDE). On each of the scatter plot, the big black thick circle corresponds to the norm at $\text{VaR}_{0.004}$ (or the 1-year return level), while the bigger grey circle corresponds to the norm at $\text{VaR}_{0.0001}$. The stars represent the real extreme points with norm greater than $\text{VaR}_{0.004}$. The little grey circles represent the simulated extreme points with norm greater than $\text{VaR}_{0.004}$ based on KDE.

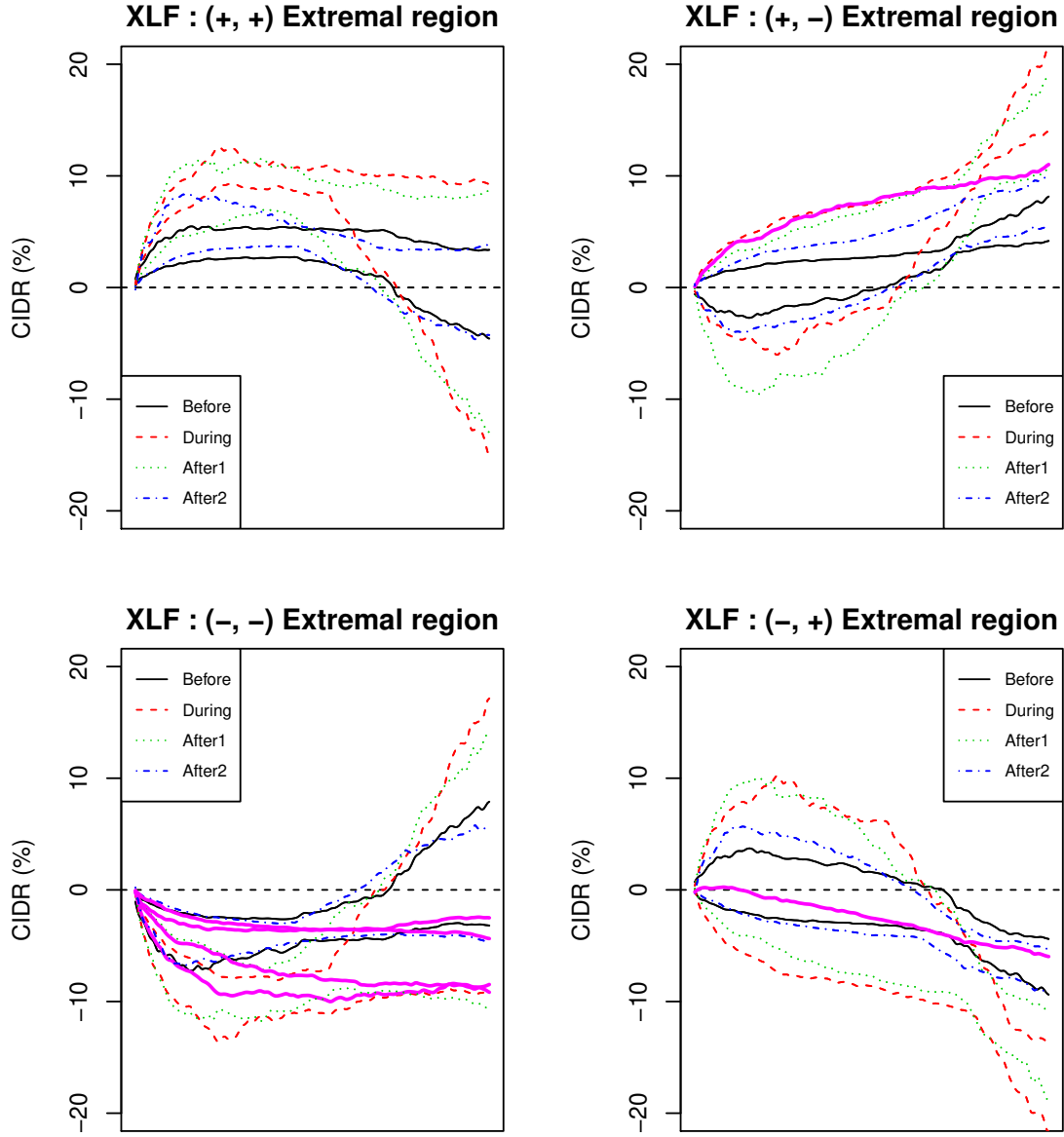


Figure 3.12: VaR Extremal regions of XLF (KDE). This is the extremal regions of CIDRs corresponding to extreme regions in different quadrants on the (ξ_1, ξ_2) plane in Figure 3.11. On each panel, the two curves in the same type, respectively, represents the lower and upper bound of the extremal region of CIDRs regarding one of the four periods under comparison. The thick solid curves represent the real extremal CIDRs with norm greater than $\text{VaR}_{0.004}$.

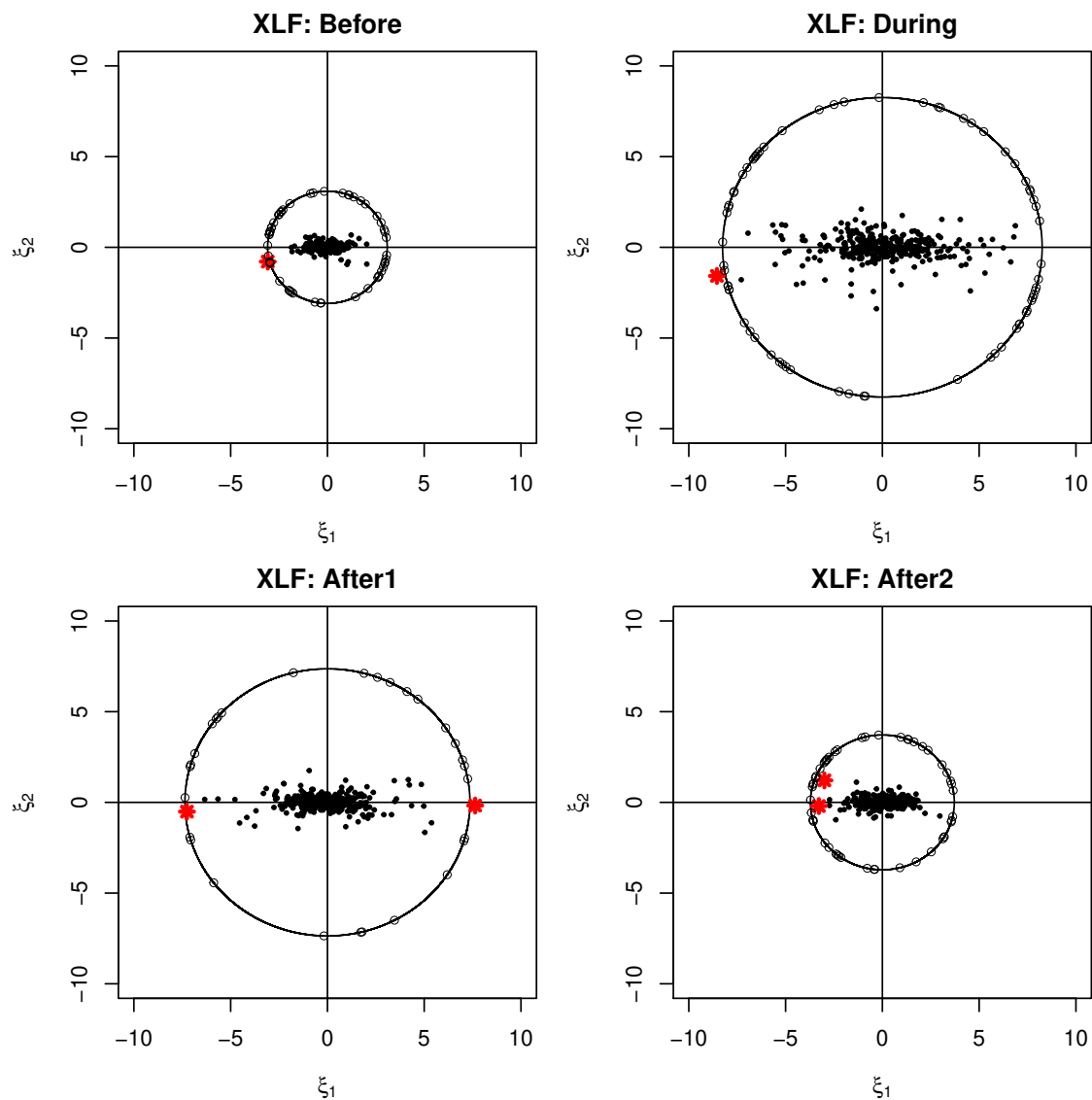


Figure 3.13: The scatter plot of XLF scores with **ES Estimates** (KDE). On each of the scatter plot, the big black thick circle corresponds to the norm at $ES_{0.004}$ (or the 1-year expected shortfall). The stars represent the real extreme points with norms greater than $VaR_{0.004}$. The little grey circles represent the simulated extreme points with norm at $ES_{0.004}$ based on KDE.

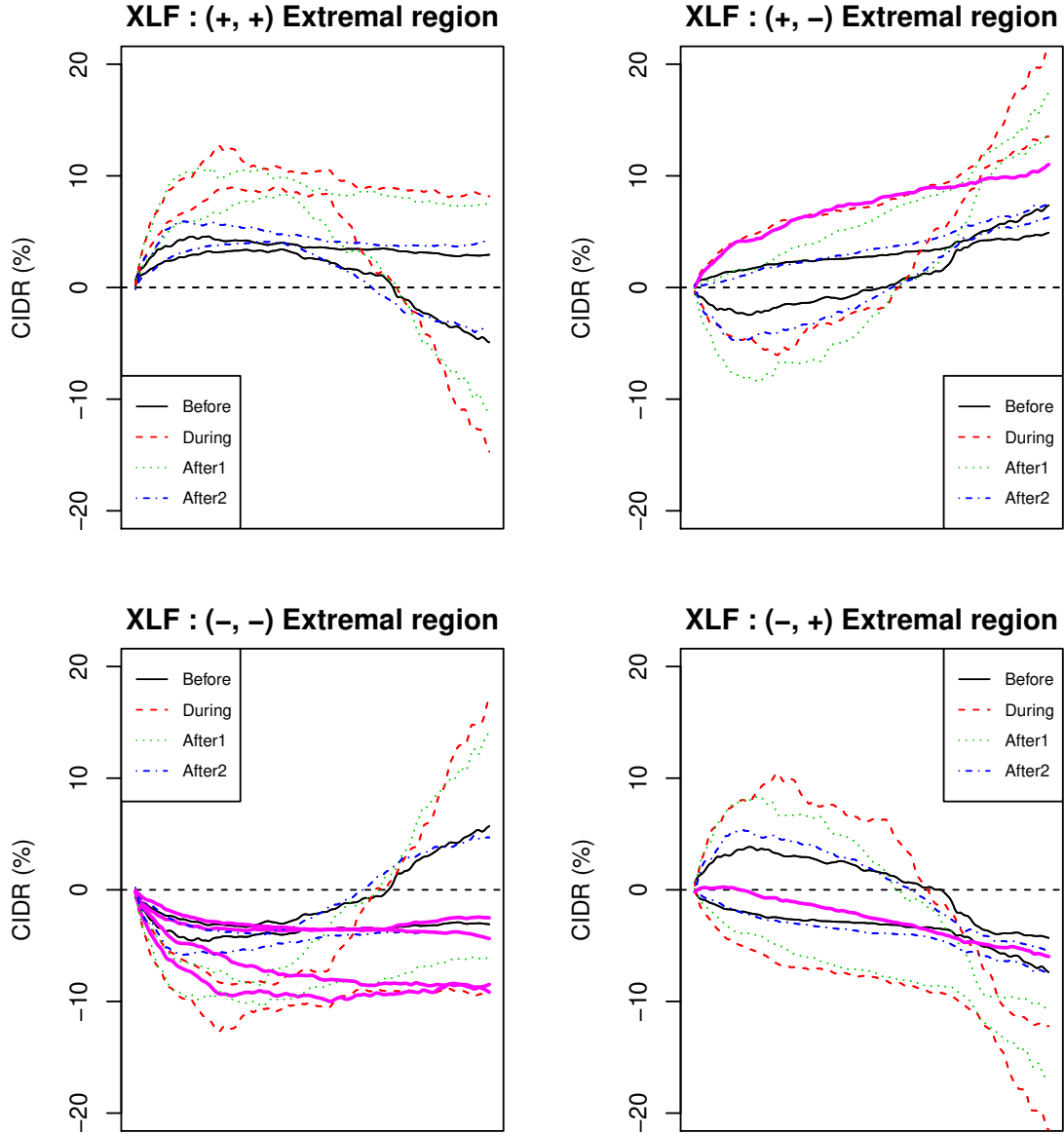


Figure 3.14: ES Extremal regions of XLF (KDE). This is the extremal regions of CIDRs corresponding to extreme regions in different quadrants on the (ξ_1, ξ_2) plane in Figure 3.13. On each panel, the two curves in the same type, respectively, represents the lower and upper bound of the extremal region of CIDRs regarding one of the four periods under comparison. The thick solid curves represent the real extremal CIDRs with norm greater than $\text{VaR}_{0.004}$.

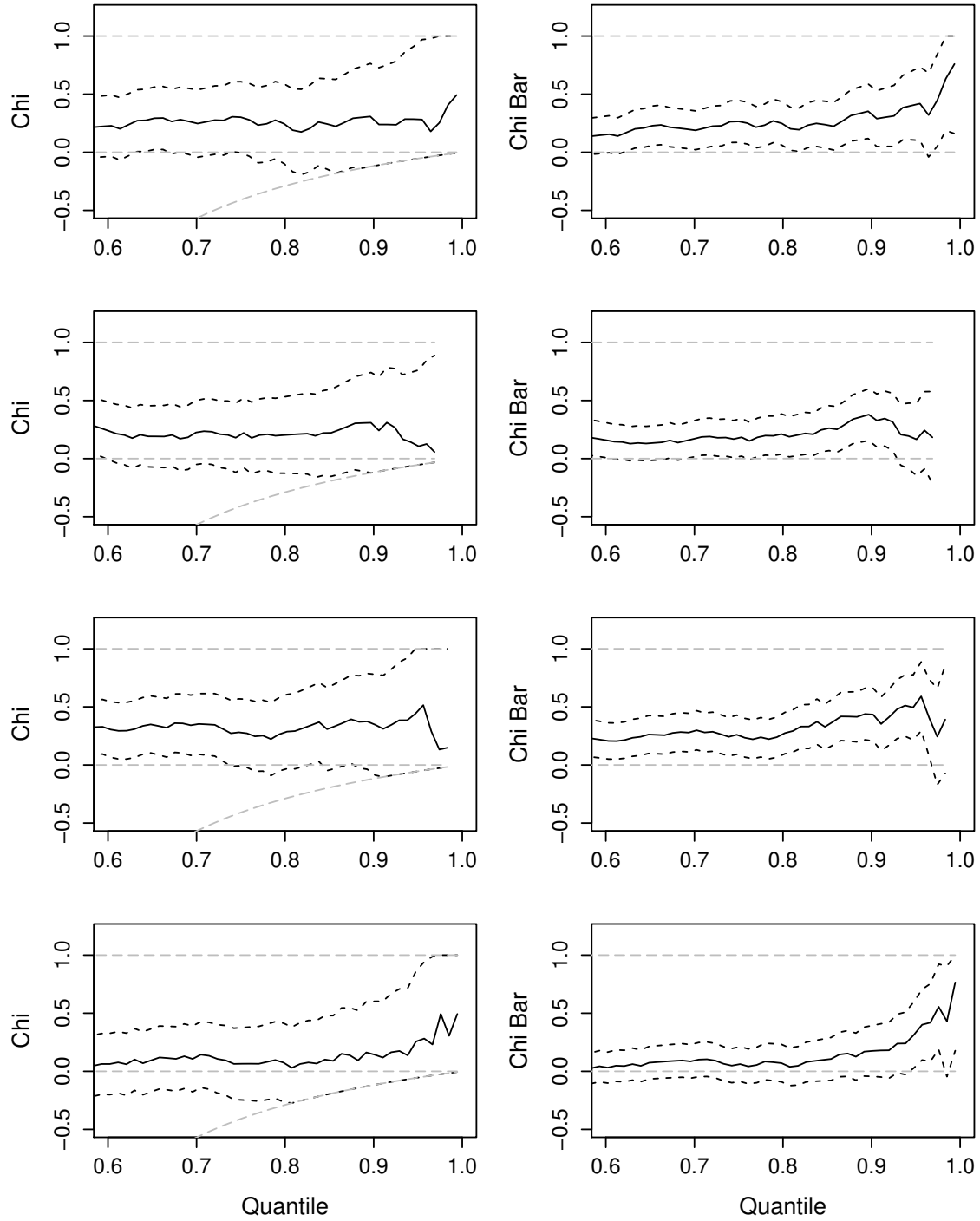


Figure 3.15: Estimates of $\chi(q)$ and $\bar{\chi}(q)$ for the squared scores of XLF. This is the estimates results of $\chi(q)$ and $\bar{\chi}(q)$ for the squared scores of the first two CIDR FPCs for XLF. Dotted lines correspond to approximate 95% confidence intervals. From top to bottom, the panels show the results for the periods “before”, “during” “after1” and “after2”.

Chapter 4

A Large-Scale Dynamic Cross-Correlation Testing

Problem

In this Chapter, we consider a large scale multiple testing problem for cross-correlations among multivariate random vectors or elements. The basic idea is inspired by [5] and an attempt is made to tailor it to address a specific hypothesis testing problem in statistical genomics and proteomics studies. The properties of the proposed method is investigated numerically at this stage. Difficulties that must be overcome are highlighted.

4.1 Introduction

Consider a two-sample problem. Let $\mathbf{X} = (X_1, \dots, X_{p_1})'$ and $\mathbf{Y} = (Y_1, \dots, Y_{p_2})'$ be two high dimensional random vectors, with dimension p_1 and p_2 respectively. It is of interest to carry out multiple correlation tests between X_i and Y_j :

$$H_{0ij} : \text{Cov}(X_i, Y_j) = 0 \quad \text{versus} \quad H_{1ij} : \text{Cov}(X_i, Y_j) \neq 0 \quad (4.1)$$

for $1 \leq i \leq p_1$ and $1 \leq j \leq p_2$, where “Cov” represents the covariance between two random variables. This is the so-called large scale multiple testing problem of cross-correlations.

Large scale multiple testing problems arise in many applications including brain connectivity analysis and gene co-expression network analysis, where thousands or millions of individual hypotheses are carried out at the same time. This brings up a particular focus on the high dimensional setting where the dimension of the vectors can be much larger than the sample size. A common goal in multiple testing is to control the false discovery rate (FDR), which is defined as the expected proportion of false positives among all rejections, i.e.

$$\text{FDR} = E[\text{False Rejections}/\text{Rejections}]$$

The difficulties of this multiple testing problem lie in (i) the construction of suitable test statistics for testing the individual hypotheses, and more importantly, in (ii) constructing a good procedure to account for the multiplicity of the test such that the overall FDR is controlled. A successful procedure must have a solid theoretical and practical control on the FDR level while maintaining a good power.

[5] propose a large scale simultaneous testing procedure for correlations in both the one-sample and two-sample settings. Their multiple testing procedure is developed in two stages. First, they construct a test statistic marginally for each individual hypothesis and show that the null distribution is asymptotically standard normal, as both the sample size and the number of individual tests go to infinity (a bootstrap method is used when the sample size is small). Second, they develop a procedure to account for the multiplicity in testing a large number of hypotheses such that the overall FDR and false discovery proportion (FDP) are controlled at least asymptotically. They also mentioned a procedure for problem (4.1) as an application of their main results. Our major objective is to tailor their procedure to solve a sophisticated multiple testing problem.

4.2 Motivation

4.2.1 Description of the data

The raw data consist of two datasets from a biological experimental study. The protein data is an 805×18 matrix, which represents the abundance profile of 805 proteins measured on 3 individual subjects repeatedly at 6 time points 0hr, 1hr, 2hr, 6hr, 12hr and 24hr. Hence the data structure can be viewed as a longitudinal type, where on each of the 3 subjects we collect data of the 805 proteins over 6 sequential time points within a day. Naturally, the 3 subjects can be viewed as a block, which give us 3 independent replicates of each protein profile. Notice that there might present substantial temporal dependence across the 6 repeated measurements for each protein within a subject.

Similarly, the genomics data is a 779×30 matrix, which stores the temporal profiles of 779 genes measured on 3 subjects at 10 time points 0hr, 20min, 40min, 60min, 90min, 120min, 240min, 360min, 12hr and 24hr. Notice that the time points for proteins do not necessarily match those for genes. All the protein and gene data have been log-transformed to make their distribution more symmetric and Gaussian-like. In summary, we have

$$\{X_{git}\} \text{ or } \mathbf{X}_i(t) = [X_{1i}(t), \dots, X_{Gi}(t)] \quad (4.2)$$

and

$$\{Y_{pis}\} \text{ or } \mathbf{Y}_i(s) = [Y_{1i}(s), \dots, Y_{Pi}(s)], \quad (4.3)$$

where $i = 1, \dots, n = 3$, $g = 1, \dots, G$ for gene and $p = 1, \dots, P$ for protein. Focusing on the first sample, we can denote by X_{git} the profile of gene g measured on subject i at time t . The vectors \mathbf{X}_i are independent. Analogous assumptions hold for the second sample. In (4.2) and (4.3), subscript “ t ” means that we view the repeated measurements as a vector, while “ (t) ” means that we treat the repeated sequential measurements as discrete samples drawn from a continuous curve.

4.2.2 The main objective

We are interested in detecting which pairs of gene and protein are correlated dynamically. We need to define a measure to quantify the similarity or correlation between each pair of gene and protein based on their dynamical profiles and threshold them after the multiplicity of the tests has been accounted for. Since our major objective is to test for the dynamic correlation among all $P \times G$ pairs simultaneously when the number of pairs is big, this problem is a large-scale multiple testing problem.

4.3 Preliminaries

The challenge of this problem comes from three major sources. First, the entries of the correlation matrix of sample statistics measuring the dynamic correlation between each protein-gene

pair are intrinsically dependent even if the original observations are independent. This attribute adds difficulty to the theory development. Second, the longitudinal data structure prevents us from using the exact procedure by [5] directly. We need to propose a new model to measure the dynamic correlation of each protein-gene pair. It is preferred that the marginal distribution of the individual test statistic under the null should be approximately normal when the sample size and dimension are both large enough. Otherwise a proper bootstrap method is needed to improve the accuracy of the approximation. Lastly, because of the temporal dependence among the repeated measurements, we have to utilize a new resampling scheme to account for the sequential dependence when designing a bootstrap procedure.

To overcome aforementioned difficulties, some new flavors are added to the conventional recipe proposed by [5]. To this end, we introduce the following essential concepts before diving into details.

4.3.1 A mixed model

Mixed models are an extension of regression models that allow for the incorporation of random effects. A common application of mixed models is the analysis of longitudinal data sets, especially, when the subjects are randomly chosen and therefore their effects can be viewed as random effects.

In the rest of this section we shall formulate the mixed model that can be used to capture the dynamic correlation between a given pair of protein and gene. For a review of mixed models, one can refer to Chapter 4 of [70].

As a way of measuring the dynamic correlation between a given pair of protein and gene, we are more interested in how the profile of each protein evolve over time can be explained by the dynamic behavior of genes. Thus in the mixed model we have Y_{pis} as the response variable denoting the profile of protein p measured on subject i at time s , where $p = 1, \dots, P = 805$ and $i = 1, \dots, n = 3$, as in Section 4.2.1. We however do not care about the real time when the protein is measured, so the index s can be simplified as $s = 1, \dots, S = 6$. The genes serve as the covariates in the model. Let us denote f_{gt} the true profiles of the gene g at time t , for $g = 1, \dots, G = 779$ and

$t = 1, \dots, T = 6$. The true gene profile f_{gt} can be estimated using the raw gene data X_{git} , since we assume

$$X_{git} = f_{gt} + \eta_{git}$$

where η_{git} is the measurement error. In practice, for each gene g , we use all 3 by 10 raw gene data points to estimate f_{g1}, \dots, f_{g6} , which are evaluated exactly at the same real time points as those available for proteins. These estimates will be denoted by \hat{f}_{gt} later on. In practice, there are many ways to estimate f_{gt} . The easiest one is just to pick the 6 real time points which match that for proteins and use the sample mean of the three replicates for each time point as the estimate \hat{f}_{gt} , or $\overline{X}_{g,t}$. This can also be done by fitting smooth curves (with either local polynomial or cubic spline methods) and then evaluate \hat{f}_{gt} at the 6 real time points on each curve.

After aligning the time points s and t at which the protein observations are available, we postulate the following marginal mixed model for each protein-gene pair (p, g) :

$$Y_{pit} = \mu_t + \alpha_{pg} + U_{pgi} + \lambda_{pg} f_{gt} + \epsilon_{pgit}, \quad (i = 1, \dots, n, t = 1, \dots, T). \quad (4.4)$$

In model (4.4), μ_t denotes the trend of all the protein profiles at time t , this term is necessary when there is a clear trend with the response over time. In practice, we can also remove this term in the estimation step by subtracting it from the response in a prior data analysis procedure. The sample mean of the responses of P proteins with all the n subjects at each time point t , $\bar{y}_{..t}$, is usually considered to be a simple and efficient estimate of μ_t . Coefficient α_{pg} is the intercept term considered to depend only on p and g . Variables U_{pg1}, \dots, U_{pgn} denote the random intercepts which can be treated as a random sample from, say, a $N(0, \sigma_{U_{pg}}^2)$ distribution for some $\sigma_{U_{pg}}^2 > 0$. There are two major merits of introducing the U_{pgi} term. First, notice that the repeated measurements within a subject is usually less variable, so it would be beneficial to utilize the *within-subject* information. Introducing U_{pgi} term as a random intercept for each subject is a good way of accounting for correlation of measurements pertaining to the same subject, or the *within-subject correlation*. Secondly, in a longitudinal study, the subjects are usually randomly sampled, so the effect of subjects is not

of interest. Random term U_{pgi} has an advantage of taking into account the randomness during sampling by introducing one single parameter, $\sigma_{U_{pg}}^2$, which is commonly referred to as a *variance component*. Slope λ_{pg} in (4.4) measures the dynamical correlation between protein p and gene g . The main goal of our work is to test

$$H_{0pg} : \lambda_{pg} = 0 \quad \text{versus} \quad H_{1pg} : \lambda_{pg} \neq 0, \quad \text{for } 1 \leq p \leq P \text{ and } 1 \leq g \leq G. \quad (4.5)$$

Lastly, we assume the random error term ϵ_{pgit} follows *i.i.d.* $N(0, \sigma_\epsilon^2)$.

Remark. An initial numerical study shows that there are very slight random effect of the subjects in our real data. So a regression model might work sufficiently well. We focus on the mixed model due to its potential theoretical value and broader scope of applications.

4.3.2 Moving block bootstrap

As a new resampling scheme, the *moving block bootstrap* (MBB) was independently proposed by [71] and [72] to be applicable to dependent data without any parametric model assumptions. In contrast to resampling individual observations by the ordinary bootstrap with *i.i.d.* observations, the MBB performs the resampling procedure within a row of formed *blocks* of consecutive observations so that the dependence structure of the original observations can be preserved within each block. Specifically, let $\mathcal{X}_n = \{X_1, \dots, X_n\}$ be the observations of a sequence of stationary random variables. Let $\mathcal{B}_i = (X_i, \dots, X_{i+l-1})$ denote the i th block of time series starting with X_i , $1 \leq i \leq N$ where $N = n - l + 1$. The block length l is an integer between 1 and n , which is usually requested to satisfy the conditions that $l \rightarrow \infty$ and $l/n \rightarrow 0$. We randomly select a suitable number (say, k) of blocks, $\mathcal{B}_1^*, \dots, \mathcal{B}_k^*$ from $\{\mathcal{B}_1, \dots, \mathcal{B}_N\}$ with replacement. Note that each of the selected blocks contains l elements and denote the elements in \mathcal{B}_i^* by $(X_{(i-1)l+1}^*, \dots, X_{il}^*)$, $i = 1, \dots, k$. Then X_1^*, \dots, X_m^* consist of the MBB sample of size $m = kl$. Typically, we choose $\lfloor n/l \rfloor$ blocks of length l with replacement from the $n - l + 1$ blocks of the observed data. When $\lfloor n/l \rfloor$ happen to be an integer then we have $m = n$, hence we can denote a random sample as $\mathcal{X}_n^* = \{X_1^*, \dots, X_n^*\}$.

Suppose $\hat{\theta}_n = T(F_n)$ is an estimation based on random sample \mathcal{X}_n , where F_n is an empirical function of random sample distribution. Then the MMB version of the estimation $\hat{\theta}_n$, denoted as θ_n^* , has the form:

$$\theta_n^* = T(F_n^*),$$

where F_n^* is the empirical distribution of MBB sample \mathcal{X}_n^* . Since each of the resampled blocks \mathcal{B}_i^* are randomly chosen from $\{\mathcal{B}_1, \dots, \mathcal{B}_N\}$ with replacement, the resampled observations $\{(X_{(i-1)l+1}^*, \dots, X_{il}^*), i = 1, \dots, k\}$ are conditionally *i.i.d* random vectors with the probability:

$$P((X_{(i-1)l+1}^*, \dots, X_{il}^*) = (X_j, \dots, X_{j+l-1})) = N^{-1},$$

for $1 \leq j \leq N$. As $l \rightarrow \infty$ with n , any finite-dimensional joint distribution of the population can be ultimately revealed from the resampled observations. That is to say, the MBB method is able to efficiently approximate features of the whole process at the population level.

As a successful variant of bootstrap, particularly for time series, MBB has acquired a broad popularity in many applications. [73] compare four different block bootstrap methods by quantifying their relative efficiency and conclude that the MBB procedure shows the lowest errors in estimation and prediction in and out of the sample overall. In practice, it is usually not preferable to directly apply MBB to original data, which requires us to create the best suitable model for the original data. Instead, we adopt the MBB procedures in resampling the series of residuals from the set of estimation models. According to the process of block replacement, it is possible some randomly selected block of residuals are not used at all, while some other one is used multiple times. An illustrative example is given in the Section 4.4.2. When the blocks of residuals are long enough, the correlation structure of the residuals should be accurately reflected through the bootstrapped ones even without the information on unknown correlations.

In Section 4.4, we apply the MBB procedure to residuals of the proposed mixed model to approximate the distribution of a studentized statistic. [74] proposed a similar idea of applying MBB for analyzing longitudinal data, which focus on a case where the number of replications over

time is large relative to the number of subjects and justify the effectiveness of MBB in the study by deriving its asymptotic properties. The proposed MBB in this paper is slightly different from [74], where we do a stratified version of MBB at and within subject level.

4.4 Large-scale multiple testing procedure

In this section, we explain our procedure for testing 4.5 in details. Theoretical results will be explored to show that the procedure asymptotically controls the FDR and FDP at pre-specified levels.

The FDR controlling procedure has been broadly used in dealing with large-scale multiple testing problems such as the genomic studies. [5] proposed a large scale multiple testing procedure for drawing inference on correlations that controls the FDR and FDP asymptotically under mild assumptions on the dependence structures. A key step is to estimate the proportion of the falsely rejected nulls by the procedure among all the true nulls at any given threshold level. Two different methods for estimating this proportion are discussed in the paper.

In the following, we consider

$$H_{0pg} : \text{protein } p \text{ is not dynamically correlated with gene } g \text{ versus} \quad (4.6)$$

$$H_{1pg} : \text{protein } p \text{ is dynamically correlated with gene } g.$$

As discussed in Section 4.3.1, mixed model 4.4 is proposed to quantify the association between each pair of gene and protein after controlling for the random effect of the random chosen subjects. So the multiple tests (4.6) amounts to (4.5).

The test statistics for the dynamical correlation of each pair can be constructed based on the studentized estimate of the slope of the mixed model, i.e.

$$T_{pg} = \frac{\hat{\lambda}_{pg}}{\widehat{sd}(\hat{\lambda}_{pg})}$$

Under proper conditions and the null hypothesis $H_{0pg} : \lambda_{pg} = 0$, as $n, T \rightarrow \infty$,

$$T_{pg} \rightarrow N(0, 1).$$

Remark. The normality of T_{pg} is not guaranteed when the sample sizes n, T are small. Another potential difficulty may arise from the estimation of the true gene profile f_{gt} . Since f_{gt} is unobservable, we have to find a proper way to estimate it from raw data. A local polynomial estimator $\hat{f}_g(t)$ is usually considered to be a good way to recover the true profile if the observed data are believed to be generated from a smooth curve $f_g(t)$ plus some disturbance errors. One can prove that for any given “ t ”,

$$\hat{f}_{gt} = f_{gt} + o_p((n \cdot T)^{-2/5}),$$

where f_{gt} represents the discrete points on the smooth curve $f_g(t)$. By incorporating the estimator \hat{f}_{gt} into the model to take the place of f_{gt} , we need to show the asymptotic normality of the newly estimated $\hat{\lambda}_{pg}$. This is also referred as a measurement error model problem.

4.4.1 FDR control procedure

Controlling the FDR or FDP is nontrivial when dependences present among the individual test statistics. [5] proposed a FDR control scheme and show that the procedure controls the FDR and FDP at the pre-specified level asymptotically. In the followings, we go through details of the FDR control procedure and make substantial adjustment to the original scheme to accommodate our situations.

The first FDR control procedure proceeds well when the normal approximation of the individual test statistics holds. We begin by showing some basic properties of the individual test statistics T_{pg} which are generally correlated. It can be proved that, under the null hypothesis H_{0pg} and certain regularity conditions,

$$\sup_{0 \leq u \leq b\sqrt{\log M}} \left| \frac{P(|T_{pg}| \geq u)}{2 - 2\Phi(u)} - 1 \right| \rightarrow 0 \quad \text{as } n \cdot T \rightarrow \infty$$

uniformly in $1 \leq g \leq G$, $1 \leq p \leq P$ and $M \leq (n \cdot T)^r$ for any $b > 0$ and $r > 0$, where Φ is the cumulative distribution function (CDF) of the standard normal distribution. In our case, $M = P \cdot G$.

Denote the set of all true null hypotheses by

$$\mathcal{H}_0 = \{(p, g) : 1 \leq p \leq P, 1 \leq g \leq G, \lambda_{pg} = 0\}$$

and $q_0 = \text{Card}(\mathcal{H}_0)$. Since the asymptotic null distribution of each test statistic T_{pg} is standard normal, it is easy to see that

$$P \left\{ \max_{(p,g) \in \mathcal{H}_0} |T_{pg}| \geq 2\sqrt{\log M} \right\} \rightarrow 0 \quad \text{as } (n, T, M) \rightarrow \infty. \quad (4.7)$$

We now develop the multiple testing procedure. Let u be the threshold level such that H_{0pg} are rejected whenever $|T_{pg}| \geq u$. Then FDP of the procedure is

$$\text{FDP}(u) := \frac{\sum_{(p,g) \in \mathcal{H}_0} I\{|T_{pg}| \geq u\}}{\max \left\{ \sum_{1 \leq p \leq P, 1 \leq g \leq G} I\{|T_{pg}| \geq u\}, 1 \right\}}. \quad (4.8)$$

An ideal threshold level for controlling the false discovery proportion at a specified level $0 < \alpha < 1$ is

$$\tilde{u} = \inf \left\{ 0 \leq u \leq 2\sqrt{\log M} : \frac{\sum_{(p,g) \in \mathcal{H}_0} I\{|T_{pg}| \geq u\}}{\max \left\{ \sum_{1 \leq p \leq P, 1 \leq g \leq G} I\{|T_{pg}| \geq u\}, 1 \right\}} \leq \alpha \right\},$$

where the constraint $0 \leq u \leq 2\sqrt{\log M}$ is used here due to the tail bound (4.7). It is easy to observe that FDP defined in (4.8) should decrease as the threshold u increases.

Threshold \tilde{u} needs to be estimated and depends on the unknown true null set \mathcal{H}_0 . The crucial step becomes to properly estimate

$$G_0(u) := \frac{1}{q_0} \sum_{(p,g) \in \mathcal{H}_0} I\{|T_{pg}| \geq u\}, \quad (4.9)$$

which is the true proportion of the falsely rejected nulls at any threshold level u . In our case P and G thus q_0 can be very large, so it is natural to use the tail of standard normal $G(u) = 2 - 2\Phi(u)$ to approximate $G_0(u)$, i.e., we have

$$\sup_{0 \leq u \leq b_M} \left| \frac{G_0(u)}{G(u)} - 1 \right| \rightarrow 0 \quad (4.10)$$

in probability as $(n, P, G) \rightarrow \infty$, where $b_M = 4\sqrt{\log M - a_M}$ and $a_M = 2 \log(\log M)$. The range $0 \leq u \leq b_M$ is nearly optimal for (4.10) to hold in the sense that a_M cannot be replaced by any constant in general. We now define the threshold used for the multiple testing procedure in the following.

Large-scale Dynamic Cross-Correlation Tests with Normal approximation (LDCCT-N) Let $0 < \alpha < 1$ and define

$$\hat{u} = \inf \left\{ 0 \leq u \leq b_M : \frac{G(u)PG}{\max \left\{ \sum_{1 \leq p \leq P, 1 \leq g \leq G} I \{ |T_{pg}| \geq u \}, 1 \right\}} \leq \alpha \right\}, \quad (4.11)$$

where $0 \leq \alpha \leq 1$ and $G(u) = 2 - 2\Phi(u)$. If \hat{u} does not exist, then set $\hat{u} = \sqrt{4 \log M}$. We reject H_{0pg} whenever $|T_{pg}| > \hat{u}$.

Remark. In the above procedure, to show that for each pair (p, g) , $T_{pg} \rightarrow N(0, 1)$, we need the sample size n , $T \rightarrow \infty$ for validate the uniform central limit theorem. Then, to show the empirical distribution function of the observed T_{pg} under nulls converges to the standard normal CDF Φ , i.e., $G_0(u) \rightarrow G(u) = 2 - 2\Phi(u)$, we need both P and G are big enough to apply the law of large number. We also require the number of the true significant alternatives is relatively small, i.e. we introduce some sparsity. Because in the above multiple testing procedure, we use total number of pairs PG as the estimate for the number $q_0 = \text{Card}(\mathcal{H}_0)$, i.e. in the “sparse” setting, we have $q_0/(PG) \approx 1$. This controls the true FDR level of the testing procedure to be close to the nominal level α .

4.4.2 FDR control procedure with MBB adjustment

As we noticed that in Section 4.4.1, the normal approximation is suitable only when the sample sizes are large. On the other hand, when the sample sizes are small, [5] suggests a bootstrap-based procedure to improve the accuracy of the approximation. Specifically, we propose a two-level stratified bootstrap procedure of the residuals from fitting data on the mixed model (4.4). First, we perform ordinary bootstrap at the subject level, as all the profiles measured on the n different subjects are independent to one another. Then, for T repeated measurements within each given subject, we apply the MBB procedure to the residuals to retain the serial dependence within a subject. Consider one protein-gene pair (p, g) . Given estimates for gene g

$$\hat{\mathbf{f}}_g = \{\hat{f}_{g1}, \dots, \hat{f}_{gT}\} \quad (4.12)$$

and the original data for protein p

$$\mathbf{y}_p = \{y_{pi1}, \dots, y_{piT}, 1 \leq i \leq n\}, \quad (4.13)$$

we fit mixed model (4.4) to $\hat{\mathbf{f}}_g$ and \mathbf{y}_p . Denote estimate $\hat{\lambda}_{pg}$ and the corresponding residuals

$$\mathbf{e}_{pg} = \{e_{pgi1}, \dots, e_{pgiT}, 1 \leq i \leq n\}. \quad (4.14)$$

Let $\mathcal{N}_n = \{1, \dots, n\}$, then at the subject level we draw $\{k_i \in \mathcal{N}_n, 1 \leq i \leq n\}$ randomly with replacement from \mathcal{N}_n . That is to say we get resamples for \mathbf{e}_{pg} as $\{e_{pgk_i1}, \dots, e_{pgk_iT}, 1 \leq i \leq n\}$. For the second stage, we apply MBB within each subject $k_i, 1 \leq i \leq n$, with the newly sampled residuals after the first bootstrap stage. Suppose each block contain l elements and the number of blocks is $k = T/l$. Implementing the MBB procedure as described in Section 4.3.2, we obtain the resampled residuals

$$\mathbf{e}_{pg}^* = \{e_{pgk_i1}^*, \dots, e_{pgk_iT}^*, 1 \leq i \leq n\}. \quad (4.15)$$

Then the bootstrapped data for protein p is $\mathbf{y}_p^* = \mathbf{y}_p - \mathbf{e}_{pg} + \mathbf{e}_{pg}^*$, or $\mathbf{y}_p^* = \{y_{pi1}^*, \dots, y_{piT}^*, 1 \leq i \leq n\}$. The entire bootstrap procedure can be applied to all the P proteins simultaneously. Thus we obtain $\mathcal{Y}^* = \{\mathbf{y}_p^*, 1 \leq p \leq P\}$ as bootstrap samples from $\mathcal{Y} = \{\mathbf{y}_p, 1 \leq p \leq P\}$. Let $\hat{\mathbf{f}} = \{\hat{\mathbf{f}}_g, 1 \leq g \leq G\}$. With \mathcal{Y}^* and $\hat{\mathbf{f}}$, we can obtain bootstrap estimate $\hat{\lambda}_{pg}^*$ and $\widehat{sd}(\hat{\lambda}_{pg}^*)$ for all the protein-gene pairs. Let

$$T_{pg}^* = \frac{\hat{\lambda}_{pg}^* - \hat{\lambda}_{pg}}{\widehat{sd}(\hat{\lambda}_{pg}^*)}. \quad (4.16)$$

For some integer $B > 0$, we replicate the above procedure B times independently and obtain $T_{pg,1}^*, \dots, T_{pg,B}^*$. Let

$$G_{n,B}^*(u) = \frac{1}{BPG} \sum_{b=1}^B \sum_{p=1}^P \sum_{g=1}^G I\{|T_{pg,b}^*| \geq u\}. \quad (4.17)$$

In the bootstrap procedure, we use the conditional distribution of $\hat{\lambda}_{pg}^* - \hat{\lambda}_{pg}$ to approximate the null distribution. Under some regularity conditions, it can be shown that, as $n, T, P, G \rightarrow \infty$,

$$\sup_{0 \leq u \leq b_M} \left| \frac{G_{n,B}^*(u)}{G_0(u)} - 1 \right| \rightarrow 0 \quad (4.18)$$

in probability. Again, we define the threshold used for the multiple testing procedure following.

Large-scale Dynamic Cross-Correlation Tests using Bootstrap (LDCCT-B). Let $0 < \alpha < 1$ and define

$$\hat{u} = \inf \left\{ 0 \leq u \leq b_M : \frac{G_{n,B}^*(u)PG}{\max \left\{ \sum_{1 \leq p \leq P, 1 \leq g \leq G} I\{|T_{pg}| \geq u\}, 1 \right\}} \leq \alpha \right\}. \quad (4.19)$$

If \hat{u} does not exist, then set $\hat{u} = \sqrt{2 \log PG}$. We reject H_{0pg} whenever $|T_{pg}| > \hat{u}$.

Remark. We give the following example to demonstrate our procedure explicitly.

Example 4.4.1. Focusing on one protein-gene pair (p, g) , as the sample size $n \times T = 3 \times 6 = 18$ is relatively small, we need to obtain $G_{n,B}^*(u)$ in (4.17). Suppose at the first bootstrap step, we obtain resampled subjects $k_1 = 1, k_2 = 2, k_3 = 2$. At the second MBB step, we obtain resampled time points $\{2, 3\}, \{3, 4\}, \{5, 6\}$ for subject $k_1 = 1$, resampled time points $\{1, 2\}, \{2, 3\}, \{5, 6\}$

for subject $k_2 = 2$ and resampled time points $\{1, 2\}, \{4, 5\}, \{4, 5\}$ for subject $k_3 = 2$. Then the bootstrap sample in (4.15) becomes

$$\begin{aligned} &e_{pg12}, e_{pg13}, e_{pg13}, e_{pg14}, e_{pg15}, e_{pg16}, \\ &e_{pg21}, e_{pg22}, e_{pg22}, e_{pg23}, e_{pg25}, e_{pg26}, \\ &e_{pg21}, e_{pg22}, e_{pg24}, e_{pg25}, e_{pg24}, e_{pg25}. \end{aligned}$$

Plugging the resampled residual data into fitted mixed model, we can easily compute a new dataset for protein p

$$y_{p11}^*, \dots, y_{p16}^*, y_{p21}^*, \dots, y_{p26}^*, y_{p31}^*, \dots, y_{p36}^*.$$

4.5 Simulation

Note that the sample size is relatively small in our gene/protein data set. To this end, in this section, we primarily report the performance of the **LDCCT-B** testing procedure by some numerical studies. The algorithm for computing the empirical FDR and power consists of the three following steps:

Step 1. The data generation procedure (DGP)

Let X_{git} denote the profile of gene g measured on i^{th} subject at t^{th} time point. For each $g \in \{1, \dots, G\}$, we generate X_{git} by

$$X_{git} = 5 + \sin(2\pi t/T \cdot g) + U_{gi} + \epsilon_{git} \quad (4.20)$$

where $U_{gi} \sim N(0, 1)$ and $\epsilon_{git} \sim N(0, 0.5)$. For each $g \in \{1, \dots, G\}$, f_{gt} was estimated from data generated by (4.20), as $\{\hat{f}_{gt}\}_{t=1, \dots, T}$. This estimation was implemented by the local polynomial method. Now let's generate the protein data. For the first q_1 proteins

$$Y_{pjt} = \hat{f}_{pt} \cdot \lambda_0 + U_{pj} + \epsilon_{pjt}, \quad p = 1, \dots, q_1 \quad (4.21)$$

and for the remaining $P - q_1$ proteins

$$Y_{pjt} = 5 \cdot \lambda_0 + U_{pj} + \epsilon_{pjt} \quad p = q_1 + 1, \dots, P, \quad (4.22)$$

where λ_0 can be a positive value (we use 0.5 or 1), $j = 1, \dots, n$ and $t = 1, \dots, T$. Similar to that for genes, we have $U_{pj} \sim N(0, 1)$, $\epsilon_{pjt} \sim N(0, 0.5)$. From the DGP (4.20), (4.21) and (4.22), we can see among all the $P \times G$ pairs of proteins and genes, only q_1 pairs of them are truly associated.

Step 2. The mixed model-based test statistic

To measure the correlation between any protein p and gene g , we postulate the following marginal mixed model for each pair (p, g)

$$Y_{pit} = \alpha_{pg} + U_{pgi} + \lambda_{pg} \hat{f}_{gt} + \epsilon_{pgit}, \quad (i = 1, \dots, n, t = 1, \dots, T) \quad (4.23)$$

as in (4.4). Notice that in our simulated data from Step 1, the number of true alternatives is q_1 and the number of true nulls is $q_0 = PG - q_1$. Let $\mathcal{H} = \{(p, g) : 1 \leq p \leq P, 1 \leq g \leq G\}$. The set of true alternative hypotheses is: $\mathcal{H}_1 = \{(k, k) : 1 \leq k \leq q_1\}$ and $\mathcal{H}_0 = \mathcal{H} / \mathcal{H}_1$. To fit the mixed model (4.23), we use the function `lme` in the R package `nlme` and obtain the coefficient $\hat{\lambda}_{pg}$ as well as its standard error.

Step 3. The FDR control procedure with MBB adjustment

In this step, we apply the bootstrap procedure and get T_{pg}^* from (4.16) and $G_{n,B}^*(u)$ from (4.17).

Let $0 < \alpha < 1$ and define

$$\hat{u} = \inf \left\{ 0 \leq u \leq b_M : \frac{G_{n,B}^*(u)PG}{\max \left\{ \sum_{1 \leq p \leq P, 1 \leq g \leq G} I \{ |T_{p,g}| \geq u \}, 1 \right\}} \leq \alpha \right\}, \quad (4.24)$$

where $M = G \cdot P$, $b_M = \sqrt{4 \log M - a_M}$ and $a_M = 2 \log(\log M)$. If \hat{u} does not exist, then let $\hat{u} = \sqrt{2 \log(PG)}$. We reject the null whenever $|T_{p,g}| \geq \hat{u}$.

Table 4.1: Empirical FDR and Power for the **LDCCT-B** procedure.

P(=G)	m	T	Empirical FDP	Empirical Power
6	5	10	0.68	1
6	5	20	0.20	1
6	5	40	0.06	1
10	5	10	0.79	0.95
10	5	20	0.56	1
10	5	40	0.08	1
20	5	10	0.88	0.92
20	5	20	0.82	0.98
20	5	40	0.57	1

The simulation results. In all simulation studies, we consider the nominal FDR level at 0.2. For all MBB procedures, we consider the length of each moving block $l = 2$ and the number of bootstraps $B = 400$. In each experiment, we consider different settings $G = P = 6, 10, 20$ and $T = 10, 20, 40$. The number of true alternatives is always defined as $q_1 = P/2$. The number of subjects is $m = 5$. We set $\lambda_0 = 1$.

We repeat all the three steps above $N = 100$ times, then calculate the average empirical FDR by

$$\text{Average} \left\{ \frac{\sum_{(p,g) \in \mathcal{H}_0} I \{ |T_{p,g}| \geq \hat{u} \}}{\max \left\{ \sum_{p,g} I \{ |T_{pg}| \geq \hat{u} \}, 1 \right\}} \right\}$$

and the average empirical powers

$$\text{Average} \left\{ \frac{\sum_{(p,g) \in \mathcal{H}_1} I \{ |T_{p,g}| \geq \hat{u} \}}{q_1} \right\},$$

where \hat{u} is an estimated threshold from the FDR control procedure.

The simulation results in terms of the empirical FDR and the empirical powers are summarized in Table 4.1. It can be seen that, for given dimensions P and G , the empirical FDP cannot be well controlled when the number of repeated measurements for each subject T is small. However, the performance of the testing procedure gets better as T increases. This is not surprising, as for large

sample size results in a better approximation of the true proportion of falsely rejected nulls among the true nulls, i.e. $G_0(u)$ in (4.9). This is a key step in estimating the unknown true null set which affects our choice of the threshold for the FDR procedure, see (4.9) and (4.10). On the other hand, as the multiplicity of the testing procedure, i.e. P and G increase, we can see the estimation of the empirical FDP is getting worse. This seems contradictory to our theoretical finding that a good approximation of $G_0(u)$ benefits from larger P and G . After a close investigation of the simulation procedure, we found that the major difficulty comes from the DGP step where we need to control the true alternative set when P and G are big while T is relatively small. The reasoning behind is perspicuous. In our DGP models, we choose $q_1 = P/2$ to reflect the “sparse” settings, whilst the proportion of true significant alternatives is $q_1/P^2 = 1/2P$. Note that q_1 gets big as P increases, and it is challenging to control the settings such that there are q_1 true alternatives as the sample size T remains small, that being said, random noise will dominate so that some true nulls and true alternatives are not distinguishable. An interesting finding from Table 4.1 implies that increasing the sample size T can be a good remedy of this problem, i.e. when the dimension P and G are big, we also need a big sample size accordingly to maintain the required FDR level. This discovery can be confirmed in the simulation study of [5] as well.

Back to the very beginning of this study, we are trying to address a sophisticated multiple testing problem when the number of individual tests is big while the sample size is relatively small (say, 3 by 6 in our real-life data). Our simulation study shows that the proposed **LDCCT-B** procedure inspired by the work of [5] could not control the FDR at a pre-specified level. A survey in this area suggests many papers working on some similar high dimensional multiple testing problems, but as far as we know, none of them really works when the sample size is really small.

Chapter 5

Summary and Future Research

Dealing with Cumulative Intraday Returns (CIDRs) is one of many appealing ways of extending traditional financial research to the functional data framework. As a specialized time series of functions, the unique strength of the CIDRs follows from their desirable properties like stationarity and quantifiable dependence structure, which are critical to deriving inferential techniques for CIDRs.

In Chapter 2, we constructed a two-sample test that can be utilized to compare if the regression coefficients, or betas, are equal for two independent populations in the context of functional data. We considered on a CIDR based functional regression model, which is used to model the dependence of the CIDRs of individual assets on the market and/or other factors. In particular, the proposed test is applied to examine the so-called “asymmetric correlations/betas” problem, which is concerned with a significant change in correlations/betas during market downturns and upturns. Our procedure is flexible and amenable to various modifications and can serve as a useful tool in the study of two sample problems in regression models. We provided both theoretical and numerical evidence that the test has the correct size and good power under some general assumptions. At last, applications of this test to a selection of financial assets using both a functional and a scalar model reveal that different return measures capture different information about the relationships between individual assets and the market factors. The proposed method is more useful for day traders, and higher frequency traders in general.

In Chapter 3, an interesting and difficult task of quantifying extremal behaviors of a time series of curves was undertaken. By representing curves with their FPC coefficients/scores, we proposed a way of depicting the extreme region of a CIDR series in terms of their magnitude and shape using the extreme value theory *peak-over-threshold* (PoT) method along with suitable nonparametric estimation methods. In our procedure, the first 2 or 3 FPC scores, which usually contribute more than 90% to the total variability, quantify the presence of a few dominating shape components of

a CIDR series, like trends or reversals. Moreover, the norm of the CIDRs can be viewed as a new measure of intraday price movements incorporating intraday volatility and changes in trend patterns. For risk analysis, the behavior of nine Sector ETFs around the 2008 financial crisis was studied to compare the traditional point-to-point returns and CIDRs. The comparison shows that there is a discrepancy in the pattern of the decay of risk measures after the crisis.

An unrelated problem of large scale multiple testing procedure was investigated in Chapter 4. The motivating science question is to detect correlations among large amount of gene-protein pairs simultaneously in a biological experimental study. An extension of the testing procedure by [5] incorporating a mixed model and moving block bootstrap method was applied to our longitudinal-like data structure and a simulation study was implemented to explore the validity of the proposed procedure. The numerical exploration shows that the required FDR level cannot be controlled when the sample size is too small. As far as we know, no current method works for this specialized high dimensional multiple test with fixed, small sample size.

Possible extensions and future work for the topics covered by this dissertation include the following directions. In Chapter 3, it might be useful to apply multivariate EVT to model the joint extremal behavior of the scores. Even though some apparent difficulties exist, it might still be possible to provide more sophisticated ways to depict extreme curves. Instead of using the PoT method, likelihood based approaches incorporating temporal dependence as well as conditional heteroskedasticity could also be feasible as a potential direction. In Chapter 4, to overcome the difficulty of small sample size or small number of repeated measurements for each subject in our particular context, it might be feasible to fit the discrete measurements of each subject with a smooth curve first so that we can draw new samples from it. This provides a possible way of having more data available for modeling.

Bibliography

- [1] R. Gabrys, L. Horváth, and P. Kokoszka. Tests for error correlation in the functional linear model. *Journal of the American Statistical Association*, 105:1113–1125, 2010.
- [2] P. Kokoszka, H. Miao, and X. Zhang. Functional dynamic factor model for intraday price curves. *Journal of Financial Econometrics*, 13:456–477, 2015.
- [3] D. Cooley and E. Thibaud. Decompositions of dependence for high-dimensional extremes. Technical report, Colorado State University, 2018.
- [4] M. Gilli and E. Këllezi. An application of extreme value theory for measuring financial risk. *Computational Economics*, 27:207–228, 2006.
- [5] T. T. Cai and W. Liu. Large-scale multiple testing of correlations. *Journal of American Statistical Association*, 111:229–240, 2016.
- [6] F. Fabozzi and J.C. Francis. Stability tests for alphas and betas over bull and bear market conditions. *Journal of Finance*, 32:1093–1099, 1977.
- [7] M.K. Kim and J.K. Zumwalt. An analysis of risk in bull and bear markets. *Journal of Financial and Quantitative Analysis*, 14:1015–1025, 1979.
- [8] J.M. Clinebell, J.R. Squires, and J.L. Stevens. Investment performance over bull and bear markets: Fabozzi and Francis revisited. *Quarterly Journal of Business and Economics*, 32:14–25, 1993.
- [9] G. Wardwood and H.M. Anderson. Does beta react to market conditions? estimates of ‘bull’ and ‘bear’ betas using a nonlinear market model with an endogenous threshold parameter. *Quantitative Finance*, 9:913–924, 2009.
- [10] R. Bhardwaj and L. Brooks. Dual betas from bull and bear markets: Reversal of the size effect. *Journal of Financial Research*, 16:269–283, 1993.

- [11] J.B. Wiggins. Betas in up and down markets. *The Financial Review*, 27:107–123, 1992.
- [12] S.W. Howton and D.R. Peterson. An examination of cross-sectional realized stock returns using a varying risk beta model. *The Financial Review*, 33:199–212, 1998.
- [13] D.J. Spiceland and J.E. Trapnell. The effect of market conditions and risk classifications on market model parameters. *The Journal of Financial Research*, 6:217–222, 1983.
- [14] W. DeBondt and R. Thaler. Further evidence on investor overreaction and stock market seasonality. *Journal of Finance*, 42:557–582, 1987.
- [15] A. Ang and J. Chen. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63:443–494, 2002.
- [16] R. Ball and S.P. Kothari. Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns. *Journal of Financial Economics*, 25:51–74, 1989.
- [17] J. Conrad, M. Gultekin, and G. Kaul. Asymmetric predictability of conditional variances. *Review of Financial Studies*, 4:597–622, 1991.
- [18] Y. H. Cho and R. F. Engle. Time-varying betas and asymmetric effects of news: Empirical analysis of blue chip stocks. Working paper, National Bureau of Economic Research, Cambridge, MA, 2000.
- [19] G. Bekaert and G. Wu. Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, 13:1–42, 2000.
- [20] Y. Hong, J. Tu, and G. Zhou. Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial Studies*, 20:1547–1581, 2007.
- [21] M. Benko, W. Härdle, and A. Kneip. Common functional principal components. *The Annals of Statistics*, 37:1–34, 2009.

- [22] L. Horváth, P. Kokoszka, and M. Reimherr. Two sample inference in functional linear models. *Canadian Journal of Statistics*, 37:571–591, 2009.
- [23] V. M. Panaretos, D. Kraus, and J. H. Maddocks. Second-order comparison of Gaussian random functions and the geometry of DNA minicircles. *Journal of the American Statistical Association*, 105:670–682, 2010.
- [24] D. Kraus and V. M. Panaretos. Dispersion operators and resistant second-order analysis of functional data. *Biometrika*, 99:813–832, 2012.
- [25] S. Fremdt, L. Horváth, P. Kokoszka, and J. Steinebach. Testing the equality of covariance operators in functional samples. *Scandinavian Journal of Statistics*, 40:138–152, 2013.
- [26] J.D. Hamilton. Oil and the macroeconomy since worldwar ii. *Journal of Political Economics*, 91:228–248, 1983.
- [27] W.E. Ferson and C.R. Harvey. Sources of risk and expected returns in global equity markets. *Journal of Banking and Finance*, 18:775–803, 1994.
- [28] C. M. Jones and G. Kaul. Oil and the stock markets. *The Journal of Finance*, 51:463–491, 1996.
- [29] S. Gogineni. Oil and the stock market: An industry level analysis. *Financial Review*, 45:995–1010, 2010.
- [30] L. Horváth, P. Kokoszka, and G. Rice. Testing stationarity of functional time series. *Journal of Econometrics*, 179:66–82, 2014.
- [31] L. Horváth and P. Kokoszka. *Inference for Functional Data with Applications*. Springer, 2012.
- [32] T. Hsing and R. Eubank. *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley, 2015.

- [33] H. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. John Wiley, 1959.
- [34] W. Sharpe. Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–442, 1964.
- [35] J. Lintner. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47:13–37, 1965.
- [36] F. Black. Capital market equilibrium with restricted borrowing. *Journal of Business*, 45:444–454, 1972.
- [37] J. O. Ramsay and B. W. Silverman. *Functional Data Analysis*. Springer, 2005.
- [38] P. Kokoszka and M. Reimherr. *Introduction to Functional Data Analysis*. CRC Press, 2017.
- [39] D. Bosq. *Linear Processes in Function Spaces*. Springer, 2000.
- [40] P. Kokoszka and M. Reimherr. Predictability of shapes of intraday price curves. *The Econometrics Journal*, 16:285–308, 2013.
- [41] D. O. Lucca and E. Moench. The pre-FOMC announcement drift. *The Journal of Finance*, 70:329–371, 2015.
- [42] X. Zhang. White noise testing and model diagnostic checking for functional time series. *Journal of Econometrics*, 194(1):76–95, 2016.
- [43] P. Kokoszka, H. Miao, and B. Zheng. Testing for asymmetry in betas of cumulative returns: Impact of the financial crisis and crude oil price. *Statistics & Risk Modeling*, 34:33–53, 2017.
- [44] P. Kokoszka, G. Rice, and H. L. Shang. Inference for the autocovariance of a functional time series under conditional heteroscedasticity. *Journal of Multivariate Analysis*, 162:35–50, 2017.
- [45] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, 1997.

- [46] A. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management*. Princeton University Press, 2005.
- [47] Y. Bensalah. Steps in applying extreme value theory to finance: A review. *Bank of Canada Working Paper*, 101:2000–20, 2000.
- [48] R. Gençay, F. Selçuk, and A. Ulugülyağci. High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics*, 33:337–356, 2003.
- [49] L. Kourouma, D. Dupre, G. Sanfilippo, and O. Taramasco. Extreme value at risk and expected shortfall during financial crisis. Technical report, HAL, 2011.
- [50] J. Danielsson and C. G. de Vries. Value-at-risk and extreme returns. *Annales d'Economie et de Statistique*, 60:239–270, 2000.
- [51] S. N. Neftci. Value-at-risk calculations, extreme events, and tail estimation. *Journal of Derivatives*, 7(3):23–38, 2000.
- [52] A. J. McNeil and R. Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, 7:271–300, 2000.
- [53] R. Gençay and F. Selçuk. Extreme value theory and value-at-risk: Relative performance in emerging markets. *International Journal of Forecasting*, 20:287–303, 2004.
- [54] C. Brooks, A. D. Clare, J. W. Dalle Molle, and G. Persaud. A comparison of extreme value theory approaches for determining value-at-risk. *Journal of Empirical Finance*, 12:339–352, 2005.
- [55] E. Brodin and C. Klüppelberg. Extreme value theory in finance. *Encyclopedia of Quantitative Risk Analysis and Assessment*, 2, 2008.

- [56] M. Rocco. Extreme value theory for finance: a survey. *Journal of Economics Surveys*, 28:82–108, 2014.
- [57] P. Giot. Market risk models for intraday data. *European Journal of Finance*, 11:309–324, 2005.
- [58] G. Dionne, P. Duchesne, and M. Pacurar. Intraday value at risk (ivar) using tick-by-tick data with application to the Toronto Stock Exchange. *Journal of Empirical Finance*, 16:777–792, 2009.
- [59] A. C. Davison and R. L. Smith. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52:393–442, 1990.
- [60] S. G. Coles, J. Heffernan, and J. A. Tawn. Dependence measures for extreme value analysis. *Extremes*, 3:5–38, 1999.
- [61] A. Aue, L. Horváth, and D. Pellatt. Functional generalized autoregressive conditional heteroskedasticity. *Journal of Time Series Analysis*, 38:3–21, 2016.
- [62] C. Ceroveckí, C. Francq, S. Hörmann, and J-M. Zakoian. Functional GARCH models: the quasi-likelihood approach and its applications. Technical report, Université libre de Bruxelles, 2018.
- [63] P. Hall, G. S. Watson, and J. Cabrera. Kernel density estimation with spherical data. *Biometrika*, 74(4):751–762, 1987.
- [64] S. I. Resnick. *Heavy-Tail Phenomena*. Springer, New York, 2007.
- [65] J. Heffernan. A directory of coefficients of tail dependence. *Extremes*, 3:3:279–290, 2000.
- [66] A. Ledford and A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187, 1996.

- [67] A. Ledford and A. Tawn. Concomitant tail behavior for extremes. *Advances in Applied Probability*, 30:197–215, 1998.
- [68] A. Stephenson. Statistics of multivariate extremes. Technical report, CRAN, 2012.
- [69] S. Poon, M. Rockinger, and J. Tawn. Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *The Review of Financial Studies*, 17:581–610, 2004.
- [70] D. Ruppert, M. P. Wand, and R. J. Carroll. *Semiparametric Regression*. Cambridge, 2003.
- [71] H. Künsch. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17:1217–1241, 1989.
- [72] R. Y. Liu and K. Singh. Efficiency and robustness in resampling. *Ann. Statist.*, 20:370–384, 1992.
- [73] B. Radovanov and A. Marcikić. A comparison of four different block bootstrap methods. *Croatian Operational Research Review*, 5:189–202, 2014.
- [74] H. Ju. Moving block bootstrap for analyzing longitudinal data. *Commun. Stat. Theory Methods*, 44(6):1130–1142, 2015.